

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2002, 18:27:05 ; Search time 4750.6 Seconds
(without alignments)
9613.790 Million cell updates/sec

Title: US-09-523-647-1

Perfect score: 2111

Sequence: 1 ttgtcaggtagtagtagaga.....catcaaatcatatcccgta 2111

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA_Main.*

- 1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq.*
- 2: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
- 3: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
- 4: /cgn2_6/ptodata/1/pna/US080_COMB.seq.*
- 5: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
- 6: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
- 7: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
- 8: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
- 9: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
- 10: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
- 11: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
- 12: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
- 13: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
- 14: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
- 15: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
- 16: /cgn2_6/ptodata/1/pna/US092_COMB.seq.*
- 17: /cgn2_6/ptodata/1/pna/US093_COMB.seq.*
- 18: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
- 19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
- 20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
- 21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
- 22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
- 23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
- 24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
- 25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
- 26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
- 27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
- 28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
- 29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
- 30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
- 31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
- 32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
- 33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
- 34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
- 35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
- 36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
- 37: /cgn2_6/ptodata/1/pna/US100_COMB.seq.*
- 38: /cgn2_6/ptodata/1/pna/US101_COMB.seq.*
- 39: /cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
- 40: /cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
- 41: /cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
- 42: /cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
- 43: /cgn2_6/ptodata/1/pna/US6004_COMB.seq.*

- 44: /cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
- 45: /cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
- 46: /cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
- 47: /cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
- 48: /cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
- 49: /cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
- 50: /cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
- 51: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
- 52: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
- 53: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
- 54: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
- 55: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
- 56: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
- 57: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
- 58: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
- 59: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
- 60: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
- 61: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
- 62: /cgn2_6/ptodata/1/pna/US6023_COMB.seq.*
- 63: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
- 64: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
- 65: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
- 66: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
- 67: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
- 68: /cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
- 69: /cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
- 70: /cgn2_6/ptodata/1/pna/US6031_COMB.seq.*
- 71: /cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
- 72: /cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
- 73: /cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
- 74: /cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
- 75: /cgn2_6/ptodata/1/pna/US6036_COMB.seq.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	IDB	ID	Description
1	2111	100.0	2111	19	US-09-523-647-1	Sequence 1, Appli
2	2106.2	99.8	1230230	18	US-09-438-185-1	Sequence 1, Appli
3	2106.2	99.8	1230230	18	US-09-438-185A-1	Sequence 378, Appli
4	1671	79.2	1671	32	US-09-841-132-378	Sequence 417, App
5	868	41.1	1659	32	US-09-841-132-417	Sequence 131, App
6	767.6	36.4	1947	37	US-10-007-693-131	Sequence 132, App
7	719.6	34.1	1278	37	US-10-007-693-132	Sequence 21, Appl
8	363.6	17.2	731	37	US-10-007-693-21	Sequence 63, Appl
9	179.4	8.5	269	18	US-09-410-568-63	Sequence 63, Appl
10	179.4	8.5	269	18	US-09-426-571-63	Sequence 63, Appl
11	179.4	8.5	269	18	US-09-454-684-63	Sequence 63, Appl
12	179.4	8.5	269	18	US-09-454-684A-63	Sequence 63, Appl
13	179.4	8.5	269	22	US-09-556-877-63	Sequence 63, Appl
14	179.4	8.5	269	22	US-09-598-419-63	Sequence 63, Appl
15	179.4	8.5	269	24	US-09-620-412A-63	Sequence 63, Appl
16	179.4	8.5	269	32	US-09-841-132-63	Sequence 63, Appl
17	120.8	5.7	800	33	US-09-868-987-2	Sequence 2, Appli
18	120.8	5.7	800	33	US-09-866-468-2	Sequence 738, App
19	48.4	2.3	414	26	US-09-666-355A-738	Sequence 285, App
20	48	2.3	6033	1	PCT-US99-26796-285	Sequence 12199, A
21	46.8	2.2	391	23	US-09-619-643-12199	Sequence 340068,
22	46.8	2.2	391	25	US-09-654-617-340068	Sequence 340068,
23	46.8	2.2	391	27	US-09-684-016-340068	Sequence 6512, Ap
24	46.8	2.2	391	53	US-60-145-485-6512	Sequence 20, Appl
25	46	2.2	48	18	US-09-461-705-20	Sequence 14, Appl
26	45.6	2.2	7218	8	US-08-466-194-14	Sequence 43, Appl
27	45	2.1	2304	26	US-09-662-254A-43	Sequence 24, Appl
28	45	2.1	50000	26	US-09-662-254-24	Sequence 24, Appl
29	45	2.1	50000	26	US-09-662-254A-24	Sequence 1549, Ap
30	44.8	2.1	543029	19	US-09-528-237A-1549	Sequence 5263, Ap
31	44.6	2.1	2000	33	US-09-887-272A-5263	

C 32 44.4 2.1 298 16 US-09-213-679A-2731 Sequence 2731, Ap
C 33 44.4 2.1 298 17 US-09-388-906A-3786 Sequence 3786, Ap
C 34 44.4 2.1 298 37 US-10-008-054-2731 Sequence 2731, Ap
C 35 44.4 2.1 298 37 US-10-023-595-2731 Sequence 2731, Ap
C 36 44.4 2.1 1141 31 US-09-806-708A-22 Sequence 22, Appl
C 37 44.2 2.1 105589 31 US-09-803-736-1477 Sequence 1477, Ap
C 38 44 2.1 255 25 US-09-654-617-114901 Sequence 114901,
C 39 44 2.1 255 27 US-09-684-016-114901 Sequence 114901,
C 40 44 2.1 256 16 US-09-244-000A-50615 Sequence 50615, A
C 41 44 2.1 256 36 US-09-978-703-50615 Sequence 50615, A
C 42 43.8 2.1 2000 33 US-09-887-272A-5263 Sequence 5263, Ap
C 43 43.6 2.1 16898 27 US-09-692-412-122 Sequence 122, App
C 44 43.6 2.1 16898 31 US-09-803-736-1102 Sequence 1102, Ap
C 45 43.6 2.1 83663 31 US-09-803-736-1481 Sequence 1481, Ap

ALIGNMENTS

RESULT 1
US-09-523-647-1
; Sequence 1, Application US/09523647
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: COHEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING
; FILE REFERENCE: 032931/0227
; CURRENT APPLICATION NUMBER: US/09/523.647
; CURRENT FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: 60/123,966
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2111
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; NAME/KEY: CDS
; LOCATION: (139)..(1809)
US-09-523-647-1

Query Match 100.0%; Score 2111; DB 19; Length 2111;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ttgatcaggttagtagagatgaattaaattcctcgactacattacattacagataataaacc 60
Db 1 ttgatcaggttagtagagatgaattaaattcctcgactacattacattacagataataaacc 60
QY 61 aaatgttgaggttagagtttacaacattctaccgattgagagagagagagagagagagag 120
Db 61 aaatgttgaggttagagtttacaacattctaccgattgagagagagagagagagagagag 120
QY 121 atgcgtagagagatccctatgtccaaactcatcagacgagtagttagcgttccttcgcgcta 180
Db 121 atgcgtagagagatccctatgtccaaactcatcagacgagtagttagcgttccttcgcgcta 180
QY 181 acgagtagatgcgtgagttgctttgcagcggggtatagagccgctgtagcagagatctctg 240
Db 181 acgagtagatgcgtgagttgctttgcagcggggtatagagccgctgtagcagagatctctg 240
QY 241 attactaagatgcgtgagttgctttgcagcggggtatagagccgctgttagcagagatctctg 300
Db 241 attactaagatgcgtgagttgctttgcagcggggtatagagccgctgttagcagagatctctg 300
QY 301 aaggttagactgtcgtgagagataaacaaccagttgaacaaagccgtgtggtgtttt 360
Db 301 aaggttagactgtcgtgagagataaacaaccagttgaacaaagccgtgtggtgtttt 360


```
QY 1021 agaagaggtaaatcaactaaocgtctgactactgttaacttactcgtggtgagacaaatgtct 1080
DB 641979 AGAAGAGGTCAATCACTAACCTTCTACTGTAAGCTTACTCGGTGGACACAAATGTCTCT 641920
QY 1081 gcaaatgaactacagttgtaataagccttgtgtacaaagtaaatatctctggtgctgat 1140
DB 641919 GCAAAATGTAACACAGTCTGTAATGAGCCTTGTGTACAAGTAATATCTCTGGTCTGAT 641860
QY 1141 tggcttacgtatgtaaaccttgagtaactctatctcagtaactcgtgaagacttg 1200
DB 641859 TGGCTTACGTATGTAACCTGTGAGTACTCTATCTCGTATCGAATCCCGGAGACTTG 641800
QY 1201 gtcttcacgtatgctgacaaagataacactccctctctggtgtacagttactgaagct 1260
DB 641799 GTTCTTCATGATGTGTCATCAAGATACACTCCCTCTCTGCTGTACAGTACTCGAAGCT 641740
QY 1261 cctgggtggagagatcgtgtaataaagctgtttggcgctattaaagaaatgtgcccagga 1320
DB 641739 CCTGGTGGAGATCTGCTGTATAAAGCTTGTGTGGCGTATTAAAGAAATGTGCCCAGGA 641680
QY 1321 gaacccctcaggtttaaaccttgatgaagctcaagttcctcgaagattcaacaatacaa 1380
DB 641679 GAACCCCTCAGTTTAAACTTGTACTGAAAGCTCAAGTCTCTGGAAGATTCACAAATCAA 641620
QY 1381 gtgcagtaactagtgagctcaactgcggaacatgtacatcttcgcgagaacaacaaca 1440
DB 641619 GTTCCAGTAACTAGTGAGTCTAACTGCGGAACATGTACATCTTGGCAGAAACAAACAA 641560
QY 1441 cattggaaggtcttgagctaccatcatgtgctgattagacacaaatgatcctctctgt 1500
DB 641559 CATTTGAAAGGTCTTGCACTACCCATATGTGCGTATTAGACACAAATGATCTCTATCTGT 641500
QY 1501 gtaggagaaataactgtctatcgtatcgtgtgaactaacctgtctctcgtgaagatact 1560
DB 641499 GTAGGAGAAATACTGTCTATCTGATCTGTGTAACTAACCGTGTCTGCTGAAGATACT 641440
QY 1561 aagctatcttbaacttgaaagttctcaaaagaacttcagccaatagcttctcaggtccca 1620
DB 641439 AAGCTATCTTAACTCTGAAGTCTTCAAAAGAATCTTCAAGCCATAGCTCTCTCAAGTCCA 641380
QY 1621 actaaaggaaacttcaggttaataccgtgtttcttcgacgctttacctaactcgggtct 1680
DB 641379 ACTAAAGGAACGATTTCAAGTAATACCGTGTGTTTTCGACGCTTTACCTAAACTCGGTCT 641320
QY 1681 aaggaatctgtagagtttctgttaccttgaaaggtattgctcccgagatgctcgcggc 1740
DB 641319 AAGGAATCTGTAGAGTTTCTCTTACCTTGAAGGTATTGCTCCCGAGATGCTCGCGC 641260
QY 1741 gaagctattcttctctgatacaactgacttcaccagttatcagacacagaaaaataccacc 1800
DB 641259 GAAGCTATTCTTCTTCTGTATACACTGACTTCACCAGTATCAGACACAGAAAATACCCAC 641200
QY 1801 gtglataaattctaagggaattatcctaagcagagcagatattccgctctgcttaggat 1860
DB 641199 GTGTATTAAATCTTAAGGAATATTCTTAAAGCAGACGATATTCCGCTCTGCTTTAGGAT 641140
QY 1861 agctttcaagaagtaaccgcttagtaccttacgtactaaagcgggttttttcttttata 1920
DB 641139 AGCTTTCAAGAAGTACCCTTTAGTACCTTACGTACTTAAGCGGTTTTTTTGTATTATA 641080
QY 1921 agctttcaatccaactcgtagagtttcttaatacaagaatattattaaagtttctgaatc 1980
DB 641079 AGCTTTCAATCCCAATCGTAGAGTTTCTTAATCAAGATATTATTTAAGTTTCTGAAATC 641020
QY 1981 ctaagatttattttaaagcccatcttttttaggtatgtaatataaaattttttaataagct 2040
DB 641019 CTAAAGATTTATTTAAAGGCCCATCTTTTATAGGTATGTAATTAATAATTAATGT 640960
QY 2041 ttctcagtgtaacctgctctcttttaggaactacactagaaacaggtatgcatcaaatc 2100
DB 640959 TTTCCTAGTGTAACTGCTCTTTTAGGAACATACACTAGAGAACGGTATGTATCAATCAATC 640900
```

```
QY 2101 tacatcccgta 2111
DB 640899 TACATCCCGTA 640889
RESULT 3
US-09-438-185A-1/c
; Sequence 1, Application US/09438185A
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1 1230230
; LENGTH: 1230230
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-438-185A-1
```

```
Query Match 99.8% Score 2106.2: DB 18: Length 1230230:
Best Local Similarity 99.9% Pred. No. 0:
Matches 2108: Conservative 0: Mismatches 3: Indels 0: Gaps 0:
QY 1 ttgacacagtgattgagagatgaattatcttcgaacttaacttaaticagataaataacc 60
DB 642999 TTGATCGGGTAGTTAGGAGATGAATTAATTCCTTGACTTACCTTAATTCAGATATAAATACCC 642940
QY 61 aaatgttgagggttaagagatttacaacacattctaccacgtatggcagagaagaaaaaaacac 120
DB 642939 AAATGTTGAGGGTAAAGAGTTTACAAAACATTTACCCGATGGCAGAGAAAATAAATATAAC 642880
QY 121 atgcgtaggagatccctctatgtccaaactcatcagacagtagttacggtcccttgcgcta 180
DB 642879 ATGCGCATAGGAGATCCCTATCTCCAAACTCATCAGACGACTACTTACGCTTCTTCCGCTA 642820
QY 181 acgagtatggcagattctcttggccagcggggatagagccgctgtacagagctcttg 240
DB 642819 ACGAGTATGGCGAGTTGCTTTGCCGCGGGGTATAGAGCCCGCTGTAGCAGAGTCTCTG 642760
QY 241 attactaagatcgtctcgttagtgcggaacacaaagccgcacctgtctctatgcagcgaag 300
DB 642759 ATTAGTAAGATCGTCGTAGTAGTCGGGAACAAAGCCAGCACCTTCTCTATGACAGCGAAG 642700
QY 301 aaggttagacttgcgttagaataaaacacacagttgaaacaaaagccggtggtctttt 360
DB 642699 AAGGTTTACACTTGTCCGTAGAAATAAACACCACTTGAACAAAAAGCCCTGCTGCTTTT 642640
QY 361 tctaataaagaatttattccctgtgaagagagcagatgcaacctgaagctcagcaa 420
DB 642639 TGTGATAAAGAAATTTTATCCCTGTGAAGAGGAGCATGTCAACCTGTAGAGGCTACGAA 642580
QY 421 gagtcttgctacgaagatgtattctctgtataaagtaaacagatgattgcacgttagaatt 480
DB 642579 GAGTCTTCTCGGAAGATTTGTATTCTCTAAAAGTAACAGATGATTTGCAACCGCTAGAAAT 642520
QY 481 tgcagtcgcttccagataacgctacgttaggattcctcttaacctattgaaaaccttgcct 540
DB 642519 TCCCASTCCGTTCCAGAAATACGCTACTGTAGGATCTCTTACCTTACCTTACCTTACCTT 642460
QY 541 ataggcaaaaagattgttggatgttggattatcacacaacagctacactgcgaagctgaa 600
```


Db 642459 ATAGGCAAAAAAGATTGTTGATGTTGTGATTACACAAACGCTACCTTGGCAAGCTGAA 642400
Qy 601 ttcgtaagcagtgatccagaacaacccctcacaagtgatgggaaattatgctgggaaatc 660
Db 642399 TTGTAAGCAGTGATCCAGAAACACTCTCAAGGTGATGGAAATATAGTCTGGAAATC 642340
Qy 661 gatcgctgggtgcaggagataaaatgcaaaattactgtatggttaaaacctcttaagaa 720
Db 642339 GATCGCTGGGTGCAGGAGATAAATGCAAAATTACTGTATGGGTAAACCTCTTAAGAA 642280
Qy 721 ggttgctgtctcacagctgactgtatgtgcttggcccagagctccgttcttatactaaa 780
Db 642279 GGTTCGTCTTACAGCTGCTACTGTATGTGCTTGCCAGAGCTCGTCTTATATAAA 642220
Qy 781 tgggtcaaccagccattgtatataagcaagaagagcactgactgtgtcttgcctaaagatgc 840
Db 642219 TCGGCTCAACGAGCAATTTGTATTAAGCAAGAAGAGGACCTGACGTGCTTGGCTTAAGATGC 642160
Qy 841 cctgatgctacaaaatcgagtagtgaacacaggatctgctattggcccgaacgtaact 900
Db 642159 CCTGTATGCTACAAATCGAAGTAGTGAACACAGGATCTGCTAATGCCCCTAAGCTAACT 642100
Qy 901 gtatataatcctgttcccgatgggtattctcatcgtcatcgtcaaaagagttctctcttt 960
Db 642099 GTAGATAATCTGTTCGGATGGCTATTTCTCATGCATCTGTCTCAAGAGTCTCTCTTT 642040
Qy 961 aacttagagacatgagacctggcgataaaaggattattacagttgagttctgcccctcaa 1020
Db 642039 AACTTAGGAGACATGAGACCTGGCGATAAAAGGTATTTACAGTTGAGTTCTGCCCTCAA 641980
Qy 1021 aagaaggtcaaatcactaacgtgtgctactgtaacttactcgtggaacacaaatgttct 1080
Db 641979 ACAAGAGTCAAACTCACTAACCTGCTACTGTAACTTACTGCGGTGGACAAATGTCT 641920
Qy 1081 gcaaatgtaactacagtggttaagttagcctgtgtacaaagtaaatatctctgggtcgtat 1140
Db 641919 GCAAAATGTAACATACAGTTGTTAATGAGCCTTGTGTACAAAGTAAATATCTCTGGTCTGAT 641860
Qy 1141 tggcttacgtatgtaaacctgtgagtaactctatctcagtcacgaatccctggagacttg 1200
Db 641859 TGGTCTTACGTATGAACCTGTGGAGTACTCTATCTCAGTATCGAATCTCGGACITG 641800
Qy 1201 gtcttcacatgctgtgacaaagatacactccctctctgggttacagtaactcgaagct 1260
Db 641799 GTTCTTCATGATGCTGTGATCCAAAGTACACTCCCTTCTGGTGTACAGTACTCGAAGCT 641740
Qy 1261 cctgggtgagagatcgtgtataaaagttgttggcggtattaaagaaatgtgccagga 1320
Db 641739 CCTGGTGGAGATCTGCTGTAATAAGTTGTTGGCGTATTAAGAAATGTGCCCAGGA 641680
Qy 1321 gaaacctccagtttaactgttagtgaagctcaagttcctggaagattccacaaatcaa 1380
Db 641679 GAAACCTCCAGTTTAACCTTGTAGTGAAGCTCAAGTTCTCGAAGATTCACAAAATCAA 641620
Qy 1381 gttgcagtaactagtgacttaactcgtggaacatgtcacctctgcgagaaacaaca 1440
Db 641619 GTTGCACTAAGTACTGACTTAACCTGCGGAACATGTATCTTGGCAGAAACAAACA 641560
Qy 1441 cattggaaggtcttgacgtaccatattgctgtatttagacacaaatgacacctatctgt 1500
Db 641559 CATTTGGAAAGGCTTTCAGCTTACCCATATGTCGCTATTAGACAAATGATCTTCTGT 641500
Qy 1501 gtaggagaaaactgtctatcgtatctgttaactaacctggttctgctgaagataact 1560
Db 641499 GTAGGAGAAAATACTGTCTATCGTATCTGTCTAATAACCTGGTGTCTGCTGAAGATACT 641440
Qy 1561 aacgtatctttaactgtgaagttctcaaaagactcagcaaatgcttcttcagggtcca 1620
Db 641439 AACGTATCTTTAATCTTGAAGTCTCAAAAGAACTTTCAGCCAAATAGCTTCTTCAGGTCCA 641380
Qy 1621 actaaaggacgatcttcaggttaataccgtgtgttttcgacgttttacctaaactcgttct 1680
Db 641379 ACTAAAGGACAGATTTTCAGGTAAATACCGTGTGTTTTCGAGCGTTTACCTAAACTCGTCTT 641320

Qy 1681 aaggaatctgtagatgtttctgttaccttgaaaggattgtctcccgagagatgctcgcggc 1740
Db 641319 AAGGAATCTGTAGAGTTTCTGTGTACCTTGAAAGGTATTGCTCCCGGAGATGCTCGCGGC 641260
Qy 1741 gaagctattcttcttctgtatcacactgacttccaccagtatcagacacagaaataccac 1800
Db 641259 GAAGCTATTCTTCTTCTGTATACACTGACTTTCACACAGTATCAGACAGAAAATACCCAC 641200
Qy 1801 gtgtattaaattcaagaataatccctaaagcagagcgatattccgctctgctttaggat 1860
Db 641199 GTGTATTAAATCTTAAGGAATTAATCCCTTAAGCAGCGCATATTCCGCTCTGCTTTAGGAT 641140
Qy 1861 agcttcaaaagacaccgcttttagtaacttactaaagcgggttttttggttttata 1920
Db 641139 AGCTTTCAAAGAAATACCGCTTTAGTACTTACGTACTTAAAGCGGTTTTTTTGTGTTTATA 641080
Qy 1921 agcttccaatccaatcgtagagtttcttaataaagatatatttaagtttctgaaatc 1980
Db 641079 AGCTCTTCAATCCAATCGTAGAGTTTCTTAAACAAGATATTAATTAAGTTTCTGAAATC 641020
Qy 1981 ctaagattttttaaagcccatctttttaggtatgtaattaaatttttaattagct 2040
Db 641019 CTAAGATTTATTTTAAAGCCCATCTTTTAGGTATGTAAATTTTAAATTTTAAATGAT 640960
Qy 2041 ttctcagtgtaacctgttctttagaaactacactagagacgagtgatgctcatcaatc 2100
Db 640959 TTTCTCTAGTGTAACTGCTCTTTTAGAACTACTACTAGGAGACGCTATGTCTCAATCAATC 640900
Qy 2101 tacatcccgta 2111
Db 640899 TACATCCCGTA 640889

RESULT 4
US-09-841-132-378
; Sequence 378, Application US/09841132
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; CURRENT FILING DATE: 2001-04-23
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841.132
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 378
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-841-132-378

Query Match 79.2%; Score 1671; DB 32; Length 1671;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1671; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 139 atgtccaaactcactcagacgagtagttacggtccttgcgttaacgagatgagcgagttgc 198
Db 1 atgtccaaactcactcagacgagtagttacggtccttgcgttaacgagatgagcgagttgc 60
Qy 199 ttgtccagcgggggtatagagcgctgttagcagagctctgattactaagatcgctcgt 258
Db 61 ttgtccagcgggggtatagagcgctgttagcagagctctgattactaagatcgctcgt 120
Qy 259 agtcggaacaaagccagcacctgttccctatgacagcgaaggttagactgttcct 318
Db 121 agtcggaacaaagccagcacctgttccctatgacagcgaaggttagactgttcct 180
Qy 319 agaatataacacacagttgaacaaaaagccgtgtgtcttttggatgataagaatttat 378

```
Db 181 agaaataaacaacacagtggaacaaaaaacggtggtgtgttttggatataaagaatttat 240
QY cctgtgaagagagagagatgtcaacotgttagaggtctagcaagagcttctgctacggaaga 438
Db 241 cctgtgaagagagagagagatgtcaacotgttagaggtctagcaagagcttctgctacggaaga 300
QY tttgtattctgttaaaagtaaacagatgattgcaacgttagaatttgcagctccgttccagaa 498
Db 301 tttgtattctgttaaaagtaaacagatgattgcaacgttagaatttgcagctccgttccagaa 360
QY taagctactgttaggactccttaccotattgaaatcttctgctataggcaaaaaagattgt 558
Db 361 taagctactgttaggactccttaccotattgaaatcttctgctataggcaaaaaagattgt 420
QY 559 gttgatgtgtgtattacacaaacagctacotcttgcgaagctgaattcgttaagcagtgatcca 618
Db 421 gttgatgtgtgtattacacaaacagctacotcttgcgaagctgaattcgttaagcagtgatcca 480
QY 619 gaaacaactctcaaaagtgaggaaattagcttggaataatcgatcgctgggtgcagga 678
Db 481 gaaacaactctcaaaagtgaggaaattagcttggaataatcgatcgctgggtgcagga 540
QY 679 gataaatgcaaaattactgtatggtgtaaaacctcttaaaaggttgcgtgttccacagct 738
Db 541 gataaatgcaaaattactgtatggtgtaaaacctcttaaaaggttgcgtgttccacagct 600
QY 739 gctactgtatgtctgccagagctcgttcttataactaaatgcgtgcgtgcgtgcgtgcgtgcgt 798
Db 601 gctactgtatgtctgccagagctcgttcttataactaaatgcgtgcgtgcgtgcgtgcgtgcgt 660
QY 799 tgtattaagcaaaagagacgtgactgtgctgctgaagatgcctgtatgctatcacaatc 858
Db 661 tgtattaagcaaaagagacgtgactgtgctgctgaagatgcctgtatgctatcacaatc 720
QY 859 gaagtgtgaacacagatctgctattgcccgttaacgtactgtatgctatcacaatc 918
Db 721 gaagtgtgaacacagatctgctattgcccgttaacgtactgtatgctatcacaatc 780
QY 919 gatggcttcttcagctatctgtgctaaagagttctctttaaactaggaagacatgaga 978
Db 781 gatggcttcttcagctatctgtgctaaagagttctctttaaactaggaagacatgaga 840
QY 979 cctggcgataaaaggtattacagttgagttgcttgcctcaaaagaggtcaaatcact 1038
Db 841 cctggcgataaaaggtattacagttgagttgcttgcctcaaaagaggtcaaatcact 900
QY 1039 aacgtgtactgttaactactgcgtgtggaacaaatgtctgcaaatgttaactacagt 1098
Db 901 aacgtgtactgttaactactgcgtgtggaacaaatgtctgcaaatgttaactacagt 960
QY 1099 gttaatgagccttgtgtacaagtaaaatctctggtgctgattggtcttactgtatgtaaa 1158
Db 961 gttaatgagccttgtgtacaagtaaaatctctggtgctgattggtcttactgtatgtaaa 1020
QY 1159 cctgtgagttactctatctcagttatcgaatcctggaagcttgcgtgctgattggtcttactgtatgtaaa 1218
Db 1021 cctgtgagttactctatctcagttatcgaatcctggaagcttgcgtgctgattggtcttactgtatgtaaa 1080
QY 1219 atccaaagatcacactccttctggtgtgtacagttactgaaagctcctggtgagagatctgc 1278
Db 1081 atccaaagatcacactccttctggtgtgtacagttactgaaagctcctggtgagagatctgc 1140
QY 1279 tgtataaaagttgttggcgattaaagaattgtgccagagagaaacccctccagttttaa 1338
Db 1141 tgtataaaagttgttggcgattaaagaattgtgccagagagaaacccctccagttttaa 1200
QY 1339 cttgtagtgaagctcaagttcctgtgaagattcacaaatcaagttgcagtaactagtgag 1398
Db 1201 cttgtagtgaagctcaagttcctgtgaagattcacaaatcaagttgcagtaactagtgag 1260
QY 1399 tctaactgcggaaatgtacattcttgcgcagaaacacacattggaaaaggtcttgcga 1458
Db 1261 tctaactgcggaaatgtacattcttgcgcagaaacacacattggaaaaggtcttgcga 1320
```

RESULT 5

US-09-841-132-417

; Sequence 417, Application US/09841132

; GENERAL INFORMATION:

; APPLICANT: Bhatia, Ajay

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Probst, Peter

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND

; FILE REFERENCE: 210121.469C8

; CURRENT APPLICATION NUMBER: US/09/841,132

; CURRENT FILING DATE: 2001-04-23

; NUMBER OF SEQ ID NOS: 599

; SOFTWARE: FastSeq for Windows Version 3.0/4.0

; SEQ ID NO 417

; LENGTH: 1659

; TYPE: DNA

; ORGANISM: Chlamydia trachomatis serovar D

US-09-841-132-417

Query Match

Best Local Similarity 41.1%; Score 868; DB 32; Length 1659;

Matches 1194; Conservative 0; Mismatches 465; Indels 27; Gaps 2;

```
QY 121 atgcgataaggagatccctatgtccaaactcatcagacagtagttacggcttcgcgcta 180
Db 1 atgcgataaggagatccctatgtcaacaaactcatcagacagtagttacggcttcgcgcta 60
QY 181 acgagtagtgcgagttgcttccagcggggtatagagccgctgtagcagagctctcg 240
Db 61 actagtgtgcgagtttatttctgtagcggggtgttagagacctatggcagagctctc 120
QY 241 attactaagatcgcgtagtagtggaacaaagccagcactgttctatgacagcgaag 300
Db 121 tctacaacgttattagcttagctgacacaa-----agcgaag 159
QY 301 aaggttagactgtccgttagaataaaacacccagttgacacaaagccgtggtgctttt 360
Db 160 gacaacacttctataaaagcaaaaaagcaaaaaacacacagcaagagactcccgta 219
QY 361 tgtgataaagaatttatccctgtgaaggggacgatgtcaacctgttagaggctcagcaa 420
Db 220 gaccgttaagaggttgcctccggttcatgag-----tctaaagctacaggacctaagag 273
QY 421 gagcttgcgaagagattgtattctgtaaaagtaaacgattgacacgtagaatt 480
|| ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
```


QY 943 caaagagttctcttcttaacttagagacatgagacctggcgataaaaggtatttaca 1002
DB 1527 CACCGTGACTGACGTTTACTCTTGGAGATATGCAACCTGGAGACACAGAACATTAAT 1468
QY 1003 gttgagttctgcccctcaaaagaggtcacaactcaactaaagttgctactgtaacttaatgc 1062
DB 1467 GTAGAGTTTGTGCGGCTTAAACGGTGGTGGCTACCAATATAGCAACGGTTTCTTACTGT 1408
QY 1063 gdtgacacaaattctctgcaaatgtaactacagttgtttaaagagccctgtgtgtaacaagta 1122
DB 1407 EGAGGACATAAAATACAGCAAGCGTAACAACCTGTGATCAATGAGCCCTGGCGTCAAGTA 1348
QY 1123 aatatctctggtgctgattggtctacgtatgtaaacctgtgagtagtctatctcagta 1182
DB 1347 ACTATTGCGAGGAGAGATGCTCTTATGTTGTAAAGCCGTGTAGAAATATGTGATCTCCGTT 1288
QY 1183 tcaaatctctgagagactggttcttctcatgattgctgataccaagatacaactccctctggt 1242
DB 1287 TCCATCTCTGGAGATCTTGTGTCGAGATGTCGCTGTGAAGACACTCTTCTCCCGGA 1228
QY 1243 gtacagttactcgaagctcctctggtgagagatctgctgtaataaagttgttggcgatt 1302
DB 1227 GTCACAGTCTTGAAGCTGCGAGGAGCTCAAAATTTCTGTGTAATAAGTAGTTTGGACGTG 1168
QY 1303 aaagaaatgtccagagagaaacccctccagtttaaaacttgtgtaaaagctcaagttcct 1362
DB 1167 AAAGAACTGAATCTCTGGAGAGCTCTTACAGTATATAAGTTCTAGTAAGAGACACAACCTCT 1108
QY 1363 ggaagattcacaaatcaagttgcagtaactagtgagcttaactgcggaacatgtacatct 1422
DB 1107 GGACAATTCACAAATTAATGTTGTGGAAGAGCTGCTCTGACTGTGGTACTTGTACTTCT 1048
QY 1423 tgcgcagaaacaaacacatctggaaggtcttgcagctaccatctatgctgattagac 1482
DB 1047 TCGCGAGAGCGACAACTTACTGGAAGAGGTGCTGCTACTCATATATGCGGTAGTAGAT 988
QY 1483 acaaatgattctatctgtgtaggagaaataactgtctatcgtatcgtgtaactaaacct 1542
DB 987 ACTTGTGACCCCTGTTGTGAGGAAATACATCTTTACCGTATTGTGTCCACCAAGA 928
QY 1543 gttctgtggaagataactaaagttatctttaaacttgaagttctcacaagaacttcagcca 1602
DB 927 GGTCTGCAAGAAATACAAATGTTCTTTAAATGCTTAAATCTCTAAAGAACTGCAACCT 868
QY 1603 atagttctcaggttccaaactaaaggaaacgatttcagggttaataccgtgttttcgagct 1662
DB 867 GTATCCTTCTCTGGACCAACTTAAAGGAACGATTACAGGCAATACAGTAGTATTGCAATCG 808
QY 1663 ttacctaaactcgtgtctaaagaaatctgttagagtttcttctgttaccttgaaggtattgct 1722
DB 807 TTACCTAGATTAGGTTCTTAAAGAAACCTGTAGAGTTTCTGTAACTTGAAGCAGTATCA 748
QY 1723 ccgagagatgctgcggaagctattcttcttcttctgatacaactgaactcaaccagtaaca 1782
DB 747 GCTGGAGATGCTGCGGGAAGCAATCTTCTTCGATACATATGATCTGCTCCAGTTTCT 688
QY 1783 gacacagaaataaccacagtgtatataaattctaa 1816
DB 687 GATACAGAGATPACACATCATTAATCTTTGA 654

RESULT 7

US-10-007-693-132/C
; Sequence 132, Application US/10007693
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157

; SEQ ID NO 132
; LENGTH: 1278
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-10-007-693-132

Query Match 34.1%; Score 719.6; DB 37; Length 1278;
Best Local Similarity 74.3%; Pred. No. 2.1e-181;
Matches 908; Conservative 0; Mismatches 314; Indels 0; Gaps 0;
QY 595 gctgaattctgaagcagttgattccagaaacaaactcctcaagtgatgggaaattagctcgg 654
DB 1278 GCAGAGTTCGTACGCGAGTGTATCCAGCGACAACCTCTACTGCTGATGCTAAAGCTTGG 1219
QY 655 aaatcgatcgctgggtcagagataaaatgaataactggtatggtgaaacactctt 714
DB 1218 AAAATTGACCGCTTAGACAAAGCGGAAAGAGTAAATTAATTAATGATGATGGTAAACCTCT 1159
QY 715 aaagaaggttgcgtctccagctgctactgattgctgtgcccagagctccctcttctat 774
DB 1158 AAAGAAGGTTGCTGCTTTACAGCTGCAACAGTATGCGCTTGTCCAGAGATCCCTTGGTT 1099
QY 775 actaaatgcggtcaaccagccatttgaataagcaagaagacctgactgctcgtccta 834
DB 1098 ACAAAATGTGGACAACCTGCTATCTGTGTTAAACAAGAGCCACAGAAATGCTGTTTG 1039
QY 835 agatgcctctgattgctacaaatcgaagtagtaacacagagatctgctattgcccgaac 894
DB 1038 CGTTGCCCACTAGTCTTACAAATTAATGTAGTGAACCAAGGACAGCAATAGCTCGTAAC 979
QY 895 gtaactgataaactcctctcccatgctattctcatgcatctgctcaagagttctc 954
DB 978 GTTGTGTTGAAATCCTGTTCCAGATGTTACGCTCACTCTCTGACACAGCTGTACTG 919
QY 955 tcttttaacttagagacatgagaaactggcgataaaaaaggtatttaccagttgagttctgc 1014
DB 918 ACGTTTACTCTTGAGATATGCAACCTGAGAGACACAGCAATTAATTAATGATGATGAT 859
QY 1015 cctcaagaagaggttcaaatcaactaactgctactgtaacttactcgtcgtgagacacaa 1074
DB 858 CCGCTTAAACGCTGGTCTGCTCTACCAATATAGCAACGGTCTTCTTACTGTGAGAGACATAA 799
QY 1075 gttctgcaaatgaactacagttgtaataagcctgtgtacaaagtaaatatctctcgt 1134
DB 798 AATACAGCAAGCGTAACAACCTGTGATCAATGAGCGCTTGCCTACAAAGTAAGTATTC 739
QY 1135 gctgattggtctcagctatgtaaaacctgtggagttactctatctcagttatcgaaatcctgga 1194
DB 738 GCAGATTGGTCTTATGTTGTAAAGCCTGTAGAATAATGATGATCTCCGTTTCCAATCTCGA 679
QY 1195 gaactgttctctcatgattgctgataccaagatacaactccctctggtgtacagtaactc 1254
DB 678 GATCTTGTGTTGGACATGCTGCTGTTGAAGACACTCTTCTCCCGAGGTACAGTCTCT 619
QY 1255 gaagctcctggtgagagatctgctgtaataaagttgttggcgatataaagaaatgctgc 1314
DB 618 GAAGCTCAGGAGCTCAAAATTTCTTAAATAAAGTAGTTTGGACTGTGCAAAAGCAAT 559
QY 1315 ccaggaaacccctccagtttaaacctgttagtgaagctcaagttcctcgtggaagattcaca 1374
DB 558 CTGGAGAGTCTCTACAGTATAAAGTTCTAGTAAGAGCACAAACTCTCTGGACAAATTCACA 499
QY 1375 aatcaagttcagtaactagttgagttcactgctggaacatgatacatctctgcgagaacaa 1434
DB 498 AATTAATGTTCTGTGAAGAGCTGCTGACTGCTGACTTGTACTTCTTCTCCGCAAGCG 439
QY 1435 acaacacattggaaggtctctgagctaccatgtagcgtattagacacaaatgactc 1494
DB 438 ACAACTTACTGGAAGAGAGTGTGCTACTCATATGTCGATGATGATGATGATGATGAT 379
QY 1495 atctgtttaggagaaaaatctctatctgattctgtgtaactaaacgtggttctctctgaa 1554

```
Db 378 GTTTGTGTAGGAGAAATACGTTTACCGGTATTGTGTCCACCAACAGAGGTTCTGCAGAA 319
Qy 1555 gatacactacgtatcttaatacttaagattctctcaaaagaacttcagccaaatagcttcttca 1614
Db 318 GATACAAATGTTCTTTAAATCTCTTAAAGAACTGCAACCTGTATCCTCTCTCT 259
Qy 1615 ggtccaaactaaaggaaagattctcaagtaataccggtgttttctcgacgctttaccctaaactc 1674
Db 258 GGACCAACTAAGGAACGATTACAGGCAATACAGTAGTAGTATTCGATTCTACCTAGATTA 199
Qy 1675 ggttctaaagaatctgaagatttctgttacccttaaaaggtattgtctccgagagatgct 1734
Db 198 GGTCTTAAAGAAACATGATAGATTCTTCTGTAACATTTGAAGACAGTATCAGCTGGAGATGCT 139
Qy 1735 cgcggcgaagctattcttctctctgtacacactgaactctccacagctatccagacacacaaaat 1794
Db 138 CGTGGGAAGCGGATTCTTTCTTCGATACATATTGACTGTTCCAGATTCTGTATACAGAGAA 79
Qy 1795 acccagctgtattaaattctaa 1816
Db 78 ACACACATCTATTAATCTTTGA 57

RESULT 8
US-10-007-693-21
; Sequence 21, Application US/10007693
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; FILE REFERENCE: 210121.515C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 21
; LENGTH: 731
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis serovar E
US-10-007-693-21

Query Match 17.2%; Score 363.6; DB 37; Length 731;
Best Local Similarity 75.8%; Pred. No. 3.9e-86;
Matches 450; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

Qy 1223 aagatacactcccttctggtgtttacagtaactcgaagctctctgtgagagatctgctga 1282
Db 4 aagacactcttctccggagtcacagctcttgagctgagctgagctcaaatctcttctgta 63

Qy 1283 ataaagtgttggcgtattaaagaaatgtgcccagagaaacccctccagtttaacttg 1342
Db 64 ataaagtgttggactgtgaaagaaactggaatctctgagagctctctacagtaataagttc 123

Qy 1343 tagtgaagctcaagtctctctggaagattccacaaatcaagttgcagtaactagtgactga 1402
Db 124 tagtaagagcaaaaactctctggacaaattccacaaataatgttgtgtgaaagctgctctg 183

Qy 1403 actgcggaacatgtacattctgcgagaaacaaacacacattggaaggtctctgcagta 1462
Db 184 actgtggtactgtactctctgcgagaaagcaactactgaaagaggtgtgctgta 243

Qy 1463 cccatagtgcgtattagacacaaatgactctctgttaggagaaataactgtctctc 1522
Db 244 ctcaatgctgctgtagatgactgtgacctgtgtgtgtaggagaaataactgtttacc 303

Qy 1523 gtatctgttaactaacoggtgtctgtgaagataactaaactatcttaactctgaagt 1582
Db 304 gtatttgtccacacagaggtctgtgcagaagatacaaatgtttctttaaagcttaaat 363

Qy 1583 tctcaagaactctcgcaaatagcttcttccaggtccaaactaaaggaacagatttcagta 1642
Db 364 tctcaagaactcgcaactgtatctctctctgtgaccaaactaaaggaacagattacagga 423
```

```
Qy 1643 atacogttatttgcagcgttttaactaaactcggttcttaaggaatctgtaagattttctg 1702
Db 424 atacagtagtattcgatttcgtttacatgattaggttcttaagaaactgtagagttttctg 483
Qy 1703 ttaccttgaaggtatttctctcccgagatgctcgcgcgaagctatttcttctctgata 1762
Db 484 taacattgaagcagtttacagctgagatgctcgttggggaagcattcttcttccgata 543
Qy 1763 cactgaactccacgtattcagacacagaaataccacagctgtattaaattcttaa 1816
Db 544 cattgaacttccagctttctgtatcacagagaatacacacattcttaattcttga 597

RESULT 9
US-09-410-568-63
; Sequence 63, Application US/09410568
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fang, Hang
; APPLICANT: Jen, Shyian
; APPLICANT: Stromberg, Erica Jean
; APPLICANT: Enghart, Susan E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.469C2
; CURRENT APPLICATION NUMBER: US/09/410,568
; CURRENT FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: fastseq for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 269
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-410-568-63

Query Match 8.5%; Score 179.4; DB 18; Length 269;
Best Local Similarity 79.2%; Pred. No. 5.4e-37;
Matches 213; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy 470 acgtagaatttgcagctccgttccagaatacgtactctaggtactcttaccctattg 529
Db 1 atgttgaatacacaagctgttcttaaatatgctacggttaggtactctctctg 60

Qy 530 aaatcttcttatagcaaaaagattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 589
Db 61 aaatctgctcaggtgaaagggatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 120

Qy 590 gcgaagctgaattcgttaagcagctccacagaacaaactcctacacagtgatgggaattag 649
Db 121 gtgaagcaggttctgtacgcagtgatccagcagcaactcctctactgctgtgtgtaggacag 180

Qy 650 tctggaaaaatcgatccctgggtgcaggagataaatgcaaaattactgtatgtgggtaaac 709
Db 181 ttggaaaaattgaccttaggacaaagcggaagagtaaaattactgtatgtgggtaaac 240

Qy 710 ctcttaaaagaaggttctgtcttccagct 738
Db 241 ctcttaaaagaaggttctgtcttccagct 269
```

```
RESULT 10
US-09-426-571-63
; Sequence 63, Application US/09426571
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fang, Hang
```



```
QY 2079 gagaacgg 2086
Db 58 GTATCTG 51

RESULT 2
US-10-027-632-211458/c
; Sequence 211458, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 211458
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-211458

Query Match
Best Local Similarity 54.8%; Score 42; DB 6; Length 684;
Matches 103; Conservative 1; Mismatches 81; Indels 3; Gaps 1;

QY 1899 aaagcggttttttttataagctcttcaatcgaatcgtagagttcttcaatcaaga 1958
Db 235 ACAGGGTTGTTGTTGTTTACCTTTTCAACAAACTCTTTGTTTCATTGATCTTTTA 176
QY 1959 tatatttaagttctgaaactcgaattatttttaaaagcccatcttttttaggtatgt 2018
Db 175 TATKTTTAAAGTCTCTATTTGTTTATGCTCTCTAACCATCATTTATTTCTTTGTCCT 116
QY 2019 aatataaatttttaattgaacttttccctagtgtaacctgtcttttaggaactacactag 2078
Db 115 ACT--AAATTTGGGTTCAGTTTGTCTCTGTTTCTTTGTTGTCCTGAGGTGCAACATGG 59
QY 2079 gagaacgg 2086
Db 58 GTATCTG 51

RESULT 3
US-09-540-209B-379/c
; Sequence 379, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540.209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 379
; LENGTH: 1227
; TYPE: DNA

; ORGANISM: B. fragilis
US-09-540-209B-379

Query Match
Best Local Similarity 57.3%; Score 41.4; DB 5; Length 1227;
Matches 75; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 1851 gcttagtagatagcttcaagaagtagccgcttttagtacctactacgtaaaagcggttttt 1910
Db 251 GCTATAGAGACGCTTATCAAAAATAAGCCGGTCAGTGGTTGAAATACAGAACGCTCACATC 192
QY 1911 ttgtttataagctcttcaatccaatcgtagagttcttcttaacaagaatattatttaagt 1970
Db 191 TGGTAAACTAGGGTTATCAATCCAACTTGAGCGAAACTTAAACGAAAGATCTTCTCTTAAAA 132
QY 1971 tctgaaatcc 1981
Db 131 AGAGGATATAC 121

RESULT 4
US-09-789-189-1188/c
; Sequence 1188, Application US/09789189
; GENERAL INFORMATION:
; APPLICANT: Lelias, Jean-Michel
; TITLE OF INVENTION: Human Polynucleotides and Polypeptides
; FILE REFERENCE: 25436/1720
; CURRENT APPLICATION NUMBER: US/09/789.189
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/183452
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 2005
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1188
; LENGTH: 285
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-789-189-1188

Query Match
Best Local Similarity 1.9%; Score 40.2; DB 5; Length 285;
Matches 111; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 1848 tctgcttttagtagatagcttcaagaagtagccgcttttagtacctactacgtaaaagcggtt 1907
Db 256 TTTTCTTTTATAGAGAGCTAAAGAGAAATAAAAAATATATATTATTGTTGGAAATTTGTA 197
QY 1908 ttttggttttataagctcttcaatccaatcgtagagttcttcttaacaagaatattatta 1967
Db 196 TTTTATTATATAAAAGGGTAGAAGTTTAAAACTGGTAAAAATTTGTTATTATTAATTA 137
QY 1968 agttctgaaatccctgaagattatttttaaaagcccatcttttaggcatgtaataaagt 2027
Db 136 TTATTATTAATTAATGATATATATTAAGAAAGAAATTTTAAATGAATAATAATAA 77
QY 2028 ttttaattaaagcttttctctagtgtaacctgtcttcttaggaactacact 2076
Db 76 TGTAAATTTAGAAATTTGTAATTTAAAAAGTATGTTTGTAGTAAAAAATTT 28

RESULT 5
US-10-027-632-116847/c
; Sequence 116847, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
```

```

; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 116847
; LENGTH: 1149
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1149)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-116847

Query Match
Best Local Similarity 1.9%; Score 39.2; DB 6; Length 1149;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 atcgggtagtagagatgaataattctgactacatttaattcaagataataaaccacaa 63
DB 807 ATATGGAATGTAGATATATATCTTCTTAACTAACTAAACGAGCTATATTTCTTA 748
QY 64 tttgagggtgaagtggttaacaaacattctaccgagtcgagagaaataaa 119
DB 747 TGTTTTGGTAATGTCATAGCCAAAATTCCTATTAATGGATCAAGAAACAATAAA 692

RESULT 6
US-10-027-632-134009/c
; Sequence 134009, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134009
; LENGTH: 566
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134009

Query Match
1.8%; Score 38.6; DB 6; Length 566;

```

```

Best Local Similarity 55.9%; Pred. No. 6.5;
Matches 71; Conservative 1; Mismatches 55; Indels 0; Gaps 0;

QY 1945 ttttaatacaagatattatttaagttcttgaaacttaagattattttaaagcccat 2004
DB 375 TTCATTATTAAATTTGTTCTGGATTTTATTAGATTGTTATCTTTATGTAACCCCTGAT 316
QY 2005 ctttttagtgatgaataataaatttttaataagcgttttcttaagtaacctgcttttt 2064
DB 315 GGTTTCAGGATATCATTTAAATTTGTTTCTTANATTTATTATTTTCCACAGACTTCCTC 256
QY 2065 agdaact 2071
DB 255 AGGAATT 249

RESULT 7
US-10-027-632-257761/c
; Sequence 257761, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 257761
; LENGTH: 629
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-257761

Query Match
1.8%; Score 38.4; DB 6; Length 629;
Best Local Similarity 52.5%; Pred. No. 7.5;
Matches 107; Conservative 0; Mismatches 96; Indels 1; Gaps 1;

QY 1835 agcgatattccgctcgtcttagtagctttcaagaagaccgcttagacct-tac 1893
DB 534 AGACTTTCTGCTCTTATTTTTCATTGTTTAAAAAATTTGTTTTTTTCTCTGCTA 475
QY 1894 gtactaaagcgggtttttgtttttataagcgtcttcaatccaatcgtagagttcttaac 1953
DB 474 ATATTAAGACTGTATATATTTTAGTTTCATTAGGTCATGACATACCTTTGCTTTCAAATA 415
QY 1954 aagatattatttaagttcttctgaaatcccaatcccaagattttttaaagcccatcttttagg 2013
DB 414 GCAAAACCTTGATCAGTTAACTGCAATTAATGACTTTGTTTAAAAAATATATATAGTGGGTAGA 355
QY 2014 tatgtaataaatttttaattaa 2037
DB 354 AATATAAGAAAAATATAAAAAATAA 331

RESULT 8
US-10-027-632-157624
; Sequence 157624, Application US/10027632

```

```
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 157624
; LENGTH: 823
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-157624

Query Match          1.8%; Score 38.4; DB 6; Length 823;
Best Local Similarity 57.5%; Pred. No. 8;
Matches 69; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 1895 tactaaagcggttttttgcataagctcttcaatccatcgtagagttcttcaatca 1954
Db 413 tagctggggatatacttttttaagtttttcaaggaaaaataatatttttcaaaa 472

QY 1955 aagatattattaaagtttctgaaatccataagatttttcaaaagcccatcttttaggt 2014
Db 473 aaaaatttttaagttttcaaaaaaacctttaaaaaaaccttaaaaaataatatttttaagt 532

RESULT 9
US-10-027-632-132469
; Sequence 132469, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 132469
; LENGTH: 898
; TYPE: DNA
; ORGANISM: Human
```

```
US-10-027-632-132469

Query Match          1.8%; Score 38.4; DB 6; Length 898;
Best Local Similarity 49.5%; Pred. No. 8.2;
Matches 96; Conservative 1; Mismatches 97; Indels 0; Gaps 0;

QY 1909 tttgtttttataagctcttcaatccatcgtagagttcttcaatcaaaagatatattttaa 1968
Db 86 tttttgggtctctattccctctcgcaataagaaaaatagttattcttcacgtcattgtgagg 145

QY 1969 gtcttgaaatccataagatttttttaaaagcccatcttttttaggtatgtataaaatt 2028
Db 146 attcaggtttatcggaattcttcaacaataagcccatcaatttttttttwaattccattt 205

QY 2029 ttttaattaaagcttttctctagtgtaacctgtcttttaggaactacacatagggagacggt 2088
Db 206 ttcaattcaggttttcaatttttctctctctggttttgagaaactataaaagggttgtaaca 265

QY 2089 tgtcatcaaatcta 2102
Db 266 tgccttaaatcta 279

RESULT 10
US-09-789-189-251
; Sequence 251, Application US/09789189
; GENERAL INFORMATION:
; APPLICANT: Lellias, Jean-Michel
; TITLE OF INVENTION: Human Polynucleotides and Polypeptides
; FILE REFERENCE: 25436/1720
; CURRENT APPLICATION NUMBER: US/09/789,189
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/183452
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 2005
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 251
; LENGTH: 359
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-789-189-251

Query Match          1.8%; Score 38; DB 5; Length 359;
Best Local Similarity 52.5%; Pred. No. 8.1;
Matches 83; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 1880 ctttagtaccttaccgtactcaaaagcggtttttttgtttataagctcttcaatccaatcgt 1939
Db 135 ctttaaccttattattttctctccacaaattatttcatattttactttaaaaattaacatat 194

QY 1940 agagttcttcaatcaaaagatatatttttaagtttctgaaatccataagatttttttaaaag 1999
Db 195 taacttttttctacatatatttttttaattacttaacattcaacttttaaaatac 254

QY 2000 cccatcttttaggtatgtaataaaaaatttttaattaa 2037
Db 255 attcctcttttaaaaaatttaattactatttttttta 292

RESULT 11
US-10-027-632-229235/c
; Sequence 229235, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
```


QY 121 YATVGSPIPIELIAIGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLWKIDRLGAG 180
DB 121 YATVGSPIPIELIAIGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLWKIDRLGAG 180
QY 181 DKCKITVWKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKOEGPDCACLRCPVCYKI 240
DB 181 DKCKITVWKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKOEGPDCACLRCPVCYKI 240
QY 241 EVVNTGSAIARNVTVDNPPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPORRGOIT 300
DB 241 EVVNTGSAIARNVTVDNPPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPORRGOIT 300
QY 301 NVATVTCGGHKCSANVTTVVNEPCVOVNIISCADMSYCKPVEYSISVSNPGDLVLDVV 360
DB 301 NVATVTCGGHKCSANVTTVVNEPCVOVNIISCADMSYCKPVEYSISVSNPGDLVLDVV 360
QY 361 IODTLPSGVTVLEAPGGEICCNKVMRIKEMCPGETLOFLKLVKAQVPGRTNOAVTSE 420
DB 361 IODTLPSGVTVLEAPGGEICCNKVMRIKEMCPGETLOFLKLVKAQVPGRTNOAVTSE 420
QY 421 SNCGTCTSCAETTHWKGLAATHMVCVLDNDPFCVGENTVYRICVTVNRGSAEDTNVSLIL 480
DB 421 SNCGTCTSCAETTHWKGLAATHMVCVLDNDPFCVGENTVYRICVTVNRGSAEDTNVSLIL 480
QY 481 KFSKELQPIASSGPTKGTISGNTVVPDLPKLGSKESVEFSVTLKGIAPGDARGEAILSS 540
DB 481 KFSKELQPIASSGPTKGTISGNTVVPDLPKLGSKESVEFSVTLKGIAPGDARGEAILSS 540
QY 541 DTLTSPVSDTENTHY 556
DB 541 DTLTSPVSDTENTHY 556

RESULT 2

US-09-523-647-2

; Sequence 2, Application US/09523647

; GENERAL INFORMATION:

; APPLICANT: MURDIN, ANDREW D.

; APPLICANT: WOMEN, RAYMOND P.

; APPLICANT: WANG, JOE

; APPLICANT: DUNN, PAMELA

; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING

; FILE REFERENCE: 032931/0227

; CURRENT APPLICATION NUMBER: 60/123,966

; PRIOR FILING DATE: 2001-06-19

; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 556

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

US-09-523-647-2

Query Match 100.0%; Score 2951; DB 19; Length 556;
Best Local Similarity 100.0%; Pred. No. 3.6e-266;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKLIRRVTVLALTSMAFCASGGIEAAVAESLITKIVASAETKPAPVPMATAKKVRLVR 60
DB 1 MSKLIRRVTVLALTSMAFCASGGIEAAVAESLITKIVASAETKPAPVPMATAKKVRLVR 60
QY 61 RNKQPEOKSRGAFCDKEFYPCIEGRCPQVEAQOESCYRLYSKVNDNDCNVEICQSVPE 120
DB 61 RNKQPEOKSRGAFCDKEFYPCIEGRCPQVEAQOESCYRLYSKVNDNDCNVEICQSVPE 120
QY 121 YATVGSPIPIELIAIGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLWKIDRLGAG 180
DB 121 YATVGSPIPIELIAIGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLWKIDRLGAG 180

QY 181 DKCKITVWKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKOEGPDCACLRCPVCYKI 240
DB 181 DKCKITVWKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKOEGPDCACLRCPVCYKI 240
QY 241 EVVNTGSAIARNVTVDNPPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPORRGOIT 300
DB 241 EVVNTGSAIARNVTVDNPPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPORRGOIT 300
QY 301 NVATVTCGGHKCSANVTTVVNEPCVOVNIISCADMSYCKPVEYSISVSNPGDLVLDVV 360
DB 301 NVATVTCGGHKCSANVTTVVNEPCVOVNIISCADMSYCKPVEYSISVSNPGDLVLDVV 360
QY 361 IODTLPSGVTVLEAPGGEICCNKVMRIKEMCPGETLOFLKLVKAQVPGRTNOAVTSE 420
DB 361 IODTLPSGVTVLEAPGGEICCNKVMRIKEMCPGETLOFLKLVKAQVPGRTNOAVTSE 420
QY 421 SNCGTCTSCAETTHWKGLAATHMVCVLDNDPFCVGENTVYRICVTVNRGSAEDTNVSLIL 480
DB 421 SNCGTCTSCAETTHWKGLAATHMVCVLDNDPFCVGENTVYRICVTVNRGSAEDTNVSLIL 480
QY 481 KFSKELQPIASSGPTKGTISGNTVVPDLPKLGSKESVEFSVTLKGIAPGDARGEAILSS 540
DB 481 KFSKELQPIASSGPTKGTISGNTVVPDLPKLGSKESVEFSVTLKGIAPGDARGEAILSS 540
QY 541 DTLTSPVSDTENTHY 556
DB 541 DTLTSPVSDTENTHY 556

RESULT 3

US-09-841-132-398

; Sequence 398, Application US/09841132

; GENERAL INFORMATION:

; APPLICANT: Bhatia, Ajay

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Probst, Peter

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND

; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

; FILE REFERENCE: 210121.469C8

; CURRENT APPLICATION NUMBER: US/09/841,132

; CURRENT FILING DATE: 2001-04-23

; NUMBER OF SEQ ID NOS: 599

; SOFTWARE: FastSeq for Windows Version 3.0/4.0

; SEQ ID NO 398

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

US-09-841-132-398

Query Match 100.0%; Score 2951; DB 22; Length 556;
Best Local Similarity 100.0%; Pred. No. 3.6e-266;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKLIRRVTVLALTSMAFCASGGIEAAVAESLITKIVASAETKPAPVPMATAKKVRLVR 60
DB 1 MSKLIRRVTVLALTSMAFCASGGIEAAVAESLITKIVASAETKPAPVPMATAKKVRLVR 60
QY 61 RNKQPEOKSRGAFCDKEFYPCIEGRCPQVEAQOESCYRLYSKVNDNDCNVEICQSVPE 120
DB 61 RNKQPEOKSRGAFCDKEFYPCIEGRCPQVEAQOESCYRLYSKVNDNDCNVEICQSVPE 120
QY 121 YATVGSPIPIELIAIGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLWKIDHLAG 180
DB 121 YATVGSPIPIELIAIGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLWKIDHLAG 180
QY 181 DKCKITVWKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKOEGPDCACLRCPVCYKI 240
DB 181 DKCKITVWKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKOEGPDCACLRCPVCYKI 240
QY 241 EVVNTGSAIARNVTVDNPPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPORRGOIT 300
DB 241 EVVNTGSAIARNVTVDNPPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPORRGOIT 300

Db 241 EVNTGSAIARNVTVDNVPDGYSHASGQVLSFNLGDMRPGDKKVFTEFCPQRRGQIT 300
Qy 301 NVATVTCGGHCKSANVTTVNNEPCVQVNIISGADMSYVCKPVEYSISVSNPGDLVLHDVV 360
Db 301 NVATVTCGGHCKSANVTTVNNEPCVQVNIISGADMSYVCKPVEYSISVSNPGDLVLHDVV 360
Qy 361 IOTLPSGVTVLEAPGGEICCNKVVWRJKEKMPGETLOFLKLVKAQVPGRTNOVAVTSE 420
Db 361 IOTLPSGVTVLEAPGGEICCNKVVWRJKEKMPGETLOFLKLVKAQVPGRTNOVAVTSE 420
Qy 421 SNGCTCTCAETTHWKGLAATHMCLVDNDPVCVNTVTRICVTRNGSAEDTNVSLIL 480
Db 421 SNGCTCTCAETTHWKGLAATHMCLVDNDPVCVNTVTRICVTRNGSAEDTNVSLIL 480
Qy 481 KFSKELOPIASSGPTKGTISGNTVVFDPALPKLGSKESVEFSVTLKGIAPGDARGEAILSS 540
Db 481 KFSKELOPIASSGPTKGTISGNTVVFDPALPKLGSKESVEFSVTLKGIAPGDARGEAILSS 540
Qy 541 DTLTSPVSDTENTHVV 556
Db 541 DTLTSPVSDTENTHVV 556

RESULT 4

US-09-438-185-559
; Sequence 559, Application US/09438185
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185
; PRIOR FILING DATE: 1999-11-11
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 559
; TYPE: PRN
; ORGANISM: Chlamydia pneumoniae
US-09-438-185-559

Query Match 100.0%; Score 2951; DB 18; Length 559;
Best Local Similarity 100.0%; Pred. No. 3.6e-266;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKLIRRVTVLTALTSMAFCASGGIEAAVAESLITKIVASAETKPAPVPMNTAKKVLVR 60
Db 4 MSKLIRRVTVLTALTSMAFCASGGIEAAVAESLITKIVASAETKPAPVPMNTAKKVLVR 63
Qy 61 RNKOPVEQKSRGAFCDKEFYPCGEGRCOPVEAQOESCYGRLYSVKVNDCNVEICQSVPE 120
Db 64 RNKOPVEQKSRGAFCDKEFYPCGEGRCOPVEAQOESCYGRLYSVKVNDCNVEICQSVPE 123
Qy 121 YATVGSYPYTEILAIAGKDCVDVITQQLPCSAEFVSSDPETTTSDGKLVWKIDRLGAG 180
Db 124 YATVGSYPYTEILAIAGKDCVDVITQQLPCSAEFVSSDPETTTSDGKLVWKIDRLGAG 183
Qy 181 DKCKITVWKPLKEGCGCFTAAATVACAPPELRSYTKGQPAICIKOEGPDCACLRCPVCYKI 240
Db 184 DKCKITVWKPLKEGCGCFTAAATVACAPPELRSYTKGQPAICIKOEGPDCACLRCPVCYKI 243
Qy 241 EVNTGSAIARNVTVDNVPDGYSHASGQVLSFNLGDMRPGDKKVFTEFCPQRRGQIT 300
Db 244 EVNTGSAIARNVTVDNVPDGYSHASGQVLSFNLGDMRPGDKKVFTEFCPQRRGQIT 303

Qy 301 NVATVTCGGHCKSANVTTVNNEPCVQVNIISGADMSYVCKPVEYSISVSNPGDLVLHDVV 360
Db 304 NVATVTCGGHCKSANVTTVNNEPCVQVNIISGADMSYVCKPVEYSISVSNPGDLVLHDVV 363
Qy 361 IOTLPSGVTVLEAPGGEICCNKVVWRJKEKMPGETLOFLKLVKAQVPGRTNOVAVTSE 420
Db 364 IOTLPSGVTVLEAPGGEICCNKVVWRJKEKMPGETLOFLKLVKAQVPGRTNOVAVTSE 423
Qy 421 SNGCTCTCAETTHWKGLAATHMCLVDNDPVCVNTVTRICVTRNGSAEDTNVSLIL 480
Db 424 SNGCTCTCAETTHWKGLAATHMCLVDNDPVCVNTVTRICVTRNGSAEDTNVSLIL 483
Qy 481 KFSKELOPIASSGPTKGTISGNTVVFDPALPKLGSKESVEFSVTLKGIAPGDARGEAILSS 540
Db 484 KFSKELOPIASSGPTKGTISGNTVVFDPALPKLGSKESVEFSVTLKGIAPGDARGEAILSS 543
Qy 541 DTLTSPVSDTENTHVV 556
Db 544 DTLTSPVSDTENTHVV 559

RESULT 5

US-09-438-185A-559
; Sequence 559, Application US/09438185A
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 559
; TYPE: PRN
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: CPN0557
US-09-438-185A-559

Query Match 100.0%; Score 2951; DB 18; Length 559;
Best Local Similarity 100.0%; Pred. No. 3.6e-266;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKLIRRVTVLTALTSMAFCASGGIEAAVAESLITKIVASAETKPAPVPMNTAKKVLVR 60
Db 4 MSKLIRRVTVLTALTSMAFCASGGIEAAVAESLITKIVASAETKPAPVPMNTAKKVLVR 63
Qy 61 RNKOPVEQKSRGAFCDKEFYPCGEGRCOPVEAQOESCYGRLYSVKVNDCNVEICQSVPE 120
Db 64 RNKOPVEQKSRGAFCDKEFYPCGEGRCOPVEAQOESCYGRLYSVKVNDCNVEICQSVPE 123
Qy 121 YATVGSYPYTEILAIAGKDCVDVITQQLPCSAEFVSSDPETTTSDGKLVWKIDRLGAG 180
Db 124 YATVGSYPYTEILAIAGKDCVDVITQQLPCSAEFVSSDPETTTSDGKLVWKIDRLGAG 183
Qy 181 DKCKITVWKPLKEGCGCFTAAATVACAPPELRSYTKGQPAICIKOEGPDCACLRCPVCYKI 240
Db 184 DKCKITVWKPLKEGCGCFTAAATVACAPPELRSYTKGQPAICIKOEGPDCACLRCPVCYKI 243
Qy 241 EVNTGSAIARNVTVDNVPDGYSHASGQVLSFNLGDMRPGDKKVFTEFCPQRRGQIT 300
Db 244 EVNTGSAIARNVTVDNVPDGYSHASGQVLSFNLGDMRPGDKKVFTEFCPQRRGQIT 303

```
QY 301 NVATVYCGGHKCSANVTTVWNEPCQVNVNIGSADMSYVCKPVEYSISVSNPGDLVLHDVV 360
Db 304 NVATVYCGGHKCSANVTTVWNEPCQVNVNIGSADMSYVCKPVEYSISVSNPGDLVLHDVV 363
QY 361 IQDTLPISGTVLEAPGGEICCNKVMYKEMCPGETLQFKLVKAQVPGFTNQAVTSE 420
Db 364 IQDTLPISGTVLEAPGGEICCNKVMYKEMCPGETLQFKLVKAQVPGFTNQAVTSE 423
QY 421 SNGCTCTCAEATTHMKGLAATHMVCVLDNDPVCVNTVYRICVTVNRGSAEDTNVSLIL 480
Db 424 SNGCTCTCAEATTHMKGLAATHMVCVLDNDPVCVNTVYRICVTVNRGSAEDTNVSLIL 483
QY 481 KFSKELOPFIASSGPTKGTISGNTVVPDLPKLGSKESVEFSVTLKGIAPGDARGEATLSS 540
Db 484 KFSKELOPFIASSGPTKGTISGNTVVPDLPKLGSKESVEFSVTLKGIAPGDARGEATLSS 543
QY 541 DTLTSPVSDTENTHVV 556
Db 544 DTLTSPVSDTENTHVV 559

RESULT 6
US-09-841-132-441
; Sequence 441, Application US/09841132
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skelky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841.132
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 441
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis serovar D
US-09-841-132-441
```

```
Query Match 73.2%; Score 2160.5; DB 22; Length 553;
Best Local Similarity 71.7%; Pred. No. 2.7e-192;
Matches 401; Conservative 75; Mismatches 68; Indels 15; Gaps 4;

QY 1 MSLKLRVTVVLTALTSMAFCFASGGIEAAVAESLITKIVASAEKTPAPVPMATAKVKRLVR 60
Db 7 MNKLIRRAVTIFAIVTSVASFASGVLETSMAESLSTNVISLADTKAK--DNTSHKSKAR 64
QY 61 RN---KOPVEOKSRGAFCDKEFYPCBEGRCQPVQEAQOESCYGRLYSVKVNDDCNVEICQS 117
Db 65 KNSKETPVNR-----KQVAPVHESKA--TGPQKDCSFGMYTVKVNDDRNVEITQA 114
QY 118 VPEYATVGSPPYPIELAIKGGKDCVDVITQOLPCEAEFVSSDPETPTSDGKLWKIDRL 177
Db 115 VPEYATVGSPPYPIELAITATGKDCVDVITQOLPCEAEFVSRSDPATPTADGKLWKIDRL 174
QY 178 GAGDKCKITVWKPLKEGCGCCTAATVACAPELRSYTKCGQPAICQKQEGPCACLRCPVC 237
Db 175 GQGEKSKITVWKPLKEGCGCCTAATVACAPELRSYTKCGQPAICQKQEGPCACLRCPVC 234
QY 238 YKIEVW--NTGSATARNVTVDNPPDGYSHASQORVLSNGLDMRPGDKVFTVEFCPOR 297
Db 235 YKININQGTATARNVVENPVDGYSHASQORVLSNGLDMRPGDKVFTVEFCPOR 294
QY 298 QITNATVTCGGHKCSANVTTVNPPCVQVNVNIGSADMSYVCKPVEYSISVSNPGDLVLH 357
Db 295 RATNATVTCGGHKCSANVTTVNPPCVQVNVNIGSADMSYVCKPVEYSISVSNPGDLVLH 354
QY 358 DVYIOTLPISGTVLEAPGGEICCNKVMYKEMCPGETLQFKLVKAQVPGFTNQAV 417
Db 355 DVYVEDTLPISGTVLEAPGGEICCNKVMYKEMCPGETLQFKLVKAQVPGFTNQAV 414
```

```
QY 418 TSNSCGTCTCAEATTHMKGLAATHMVCVLDNDPVCVNTVYRICVTVNRGSAEDTNVS 477
Db 415 KSCSDCGTCTCAEATTHMKGLAATHMVCVLDNDPVCVNTVYRICVTVNRGSAEDTNVS 474
QY 478 LILKFSKELOPFIASSGPTKGTISGNTVVPDLPKLGSKESVEFSVTLKGIAPGDARGEAT 537
Db 475 LMLKFSKELOPFIASSGPTKGTITGNTVVPDLPKLGSKESVEFSVTLKAVSACDARGEAT 534
QY 538 LSSDTLTPVSDTENTHVV 556
Db 535 LSSDTLTPVSDTENTHVV 553

RESULT 7
US-09-201-228A-1036
; Sequence 1036, Application US/09201228A
; GENERAL INFORMATION:
; APPLICANT: Griffais, Remy
; APPLICANT: Holseth, Susan K.
; APPLICANT: Zagursky, Robert John
; APPLICANT: Metcalf, Benjamin J.
; APPLICANT: Peek, Joel A.
; APPLICANT: Sankaran, Banumathi
; APPLICANT: Fletcher, Leah Diane
; TITLE OF INVENTION: CHLAMYDIA TRACHOMATIS GENOMIC SEQUENCE
; TITLE OF INVENTION: AND POLYPEPTIDES, FRAGMENTS THEREOF AND USES THEREOF, IN
; FILE REFERENCE: 9710-0004-999
; CURRENT APPLICATION NUMBER: US/09/201.228A
; CURRENT FILING DATE: 1998-11-30
; PRIOR APPLICATION NUMBER: US 60/107,077
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: FR 97-16034
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: FR 97-15041
; PRIOR FILING DATE: 1997-11-28
; NUMBER OF SEQ ID NOS: 5981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1036
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-09-201-228A-1036
```

```
Query Match 71.3%; Score 2104.5; DB 16; Length 557;
Best Local Similarity 70.2%; Pred. No. 4.7e-187;
Matches 393; Conservative 76; Mismatches 73; Indels 19; Gaps 6;

QY 1 MSLKLRVTVVLTALTSMAFCFASGGIEAAVAESLITKIVASAEKTPAPVPMATAKVKRLVR 60
Db 7 MNKLIRRAVTIFAIVTSVASFASGVLETSMAESLSTNVISLADTKAK--DNTSHKSKAR 64
QY 61 RN---KOPVEOKSRGAFCDKEFYPCBEGRCQPVQEAQOESCYGRLYSVKVNDDCNVEICQS 117
Db 65 KNSKETPVNR-----KQVAPVHESKA--TGPQKDCSFGMYTVKVNDDRNVEITQA 114
QY 118 VPEYATVGSPPYPIELAIKGGKDCVDVITQOLPCEAEFVSSDPETPTSDGKLWKIDRL 177
Db 115 VPEYATVGSPPYPIELAITATGKDCVDVITQOLPCEAEFVSRSDPATPTADGKLWKIDRL 174
QY 178 GAGDKCKITVWKPLKEGCGCCTAATVACAPELRSYTKCGQPAICQKQEGPCACLRCPVC 237
Db 175 GQGEKSKITVWKPLKEGCGCCTAATVACAPELRSYTKCGQPAICQKQEGPCACLRCPVC 234
QY 238 YKIEVW--NTGSATARNVTVDNPPDGYSHASQORVLSNGLDMRPGDKVFTVEFCPOR 295
Db 235 YKININQGTATARNVVENPVDGYSHASQORVLSNGLDMRPGDKVFTVEFCPOR 294
QY 298 QITNATVTCGGHKCSANVTTVNPPCVQVNVNIGSADMSYVCKPVEYSISVSNPGDLVLH 355
Db 295 RATNATVTCGGHKCSANVTTVNPPCVQVNVNIGSADMSYVCKPVEYSISVSNPGDLVLH 354
```

Db 295 RGRATNTAMVYCGGHKNTASVTTVINPCVQVSIAGADWSYVCKPVEYVISVNSPGDLY 354
Qy 356 LHDVVIOQLPSTGVTVLEAPGGEICCNKVVWRIKEMCPGETLQFLKLVK--AQVPGRFN 413
Db 355 LRDVVVKTLPSTGVTVLEAAGASICNKKVWTVKELNPGESLOYKVLVARGAQTPOGFN 414
Qy 414 QVAVTSNCGTCTSCAETTHMKGLAATHMCVLDTDNDPICVGENTVYRICVNRGSAED 473
Db 415 NVVVKSCDCGTCTSCAETTVKGVAAATHMCVVDTCDPVCVGENTVYRICVNRGSAED 474
Qy 474 TNVSLILKFSKLOPSTASGPTKGTISGNTVVFDALPKLGSKEVSEFVTLKGIAPGDAR 533
Db 475 TNVSLMLKFSKLOPSTASGPTKGTITGNTVVFDLPLRLSKETVEFVTLKAVSAGDAR 534
Qy 534 GEALSSDLSLTPSVSTENTHY 556
Db 535 GEALSSDLSLTPSVSTENTHY 557

RESULT 8
US-10-020-269-38
; Sequence 38, Application US/10020269
; GENERAL INFORMATION:
; APPLICANT: Mitchell, William M.
; TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE
; FILE REFERENCE: VDB98-01
; CURRENT APPLICATION NUMBER: US/10/020,269
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/025,596
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/023,921
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-08-14
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 38
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-020-269-38

Query Match 4.8%; Score 141; DB 24; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 60 RHNKQPVQKSGAFCDKEFPCEE 84
Db 1 RHNKQPVQKSGAFCDKEFPCEE 25

RESULT 9
US-09-902-540-10065
; Sequence 10065, Application US/09902540
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10065
; LENGTH: 902
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10065

Query Match 4.5%; Score 132; DB 23; Length 902;
Best Local Similarity 22.8%; Pred. No. 0.022;
Matches 104; Conservative 40; Mismatches 161; Indels 152; Gaps 23;
Qy 103 SVKVNDDC--NVEICQ-----SVPEYATVGPYPTEILAIKGGKDCVDDVVTIQLPCEA 153
Db 517 TVTVDGAGNASTCQFTVTVRDTTAP---TIVCPAEVVEAMGPAGAVAGFV---LPQAT 570
Qy 154 ERFVSSDPETT--PTSDGKL-----VWKIDRLGAGDKCKITVWVKPLKGGCCCTAATV 203
Db 571 DAVTASPOVTASPVSGSTFMLGSTPVTLTATDDAGNVGSCQFLTVRD-----TAAPV 623
Qy 204 CACPRLSRYTRCGQPAICIKQSGPDCAKCPVYKIEVNTGSAIARNVTVDNBPQGY 263
Db 624 LTCPE-----EVIVEATGAGATVSPAGQAVDVV-----SOVQPVY 660
Qy 264 SHASGORVLSFNLGDMRPGDKKVFTEFCFQRRGQTNVATVTCGHHKCSANVTTVVNE 323
Db 661 SHASG---ADFLG;-----GTLSV--TATDAAGNAACQFTVTVRD 697
Qy 324 P-----CVQ--VNISGADMSYCKPVEYSISVSNPGLVLDVHVVIOQLPSTGVTVLEAPG 376
Db 698 TTPPELGCPQDVLAEAD---TQCAVVTFAGVQPRDLVTHDPSIAFSOTSG-----745
Qy 377 GEICCNKVVWRIKEMCPGETLQFLKLVKQVPGRF---TNOAVTSESNCGTCTCAETT 433
Db 746 -----SRFPLGTTAVAVTASDAAGNEASCRAV 773
Qy 434 THWKGUATHMCVLD---TNDPTCVGENTVYRICVNRGSAEDTNVSLILKFSKLOPI 489
Db 774 TVDPTAPQVCMPSDVSVETQDP---EGTVLSYA---PASAVDGVSSVTVVAYS-----L 821
Qy 490 ASSGPTKGTISGNT--VFDALPKLGSKEVSEFVSTLK 525
Db 822 ASGG---HFASGTTVPVVVTTATDTAGNAACQSFVLSVR 855
RESULT 10
US-07-808-458-2
; Sequence 2, Application US/07808458
; GENERAL INFORMATION:
; APPLICANT: Fleming, Robert J.
; TITLE OF INVENTION: Nucleotide And Protein Sequences Of The
; TITLE OF INVENTION: Scirrate Gene And Methods Based Thereon
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/808,458
; FILING DATE: 19911211
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-005
; TELEPHONE: 212 790-9090
; TELEFAX: 212 8698864/9741
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1404 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-07-808-458-2

Query Match 4.4%; Score 130.5; DB 3; Length 1404;
Best Local Similarity 20.5%; Pred. No. 0.058;
Matches 124; Conservative 57; Mismatches 208; Indels 217; Gaps 32;
Qy 14 LTMASCFASGGIEAAVAESLITKIVASAETKPAVPMTAKKVLRRNKPVEQKSRGA 73
Db 543 LTTTATIGNSLSTALLAALTSAVASTSLAIGPC-INAKECR-----NQP-----GS 590
Qy 74 FDCKEFYPCBEGRCQVPEAQOESCYSYGRLYSVKVNDCNVEICQSPVEYATVGSPIEL 133
Db 591 FA-----CICKEG-----WGGVTCAENLDDC-VGOCRN----- 617
Qy 134 AIGKKDQVDVITQOLPCEAEFVSDPE-----TPTSDG-----KLVWK-----IDRLG- 178
Db 618 ---GATCIDLVNDYRCACASGFTGRDCTDIDECATSPCRNGGECVDMVGKFCNICPLGY 674
Qy 179 AGDKC---KITVWVKPLKGGCFTA--ATVCAPELRSYTKCQO-PAIC----- 221
Db 675 SGLCEAEKENCPTSPCLEGHCLNTPGGYCHCPDRAGKHCEQLRPLCSQPPNEGCPA 734
Qy 222 -----IKQEGPDC-----ACLRCPV---CYKIEVVNTGSAIA 250
Db 735 NVSLATSATTTTTTTTATTTTRKMAKPSGLPCSGHSGCEMSDVGTFC-KCHVGHGTTCFE 793
Qy 251 RNVTVDNFVP-----DG-----YSHASQORVLS-----FNLGDMRPG-- 282
Db 794 HNLNECSPNCRNGGICLDGDDFTCECMGSGWTGKRCERATGCGYAGOCNGGTCMPGAP 853
Qy 283 DKV-----FTVEFCPQRRGQITNVATVTCGGHKCSANVTTVVNEPCVQVNSGA 333
Db 854 DKALOPHCRCAPGWTGLFCAE-----AIDOCROPCCHNGGT-----CE----SCA 894
Qy 334 DW-SYVCKPVEYSISVSNPGDLVLDVVIQDTPSGVTVLEAPGECICCNKVVVRKEMC 392
Db 895 GWFRVC-----AQGSGPDCRINNECSPPQCGGATCIGDIGYS-----IC 939
Qy 393 PGETLQKLVVKAQVGRFTNOVATVSESNGCTCTCAET-----TTHWKGAAATH 443
Db 940 P-----PGRHLRCEILLSDPKSACONASNTISPYTALNRSQNWLDIALTG 985
Qy 504 VVFDAL 509
Db 1036 VCVPAL 1041

RESULT 11
US-09-195-524-2
; Sequence 2, Application US/09195524
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowitz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; NUMBER OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/195,524
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/611,729
; FILING DATE: 06-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1404 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-09-195-524-2

Query Match 4.4%; Score 130.5; DB 15; Length 1404;
Best Local Similarity 20.5%; Pred. No. 0.058;
Matches 124; Conservative 57; Mismatches 208; Indels 217; Gaps 32;
Qy 14 LTMASCFASGGIEAAVAESLITKIVASAETKPAVPMTAKKVLRRNKPVEQKSRGA 73
Db 543 LTTTATIGNSLSTALLAALTSAVASTSLAIGPC-INAKECR-----NQP-----GS 590
Qy 74 FDCKEFYPCBEGRCQVPEAQOESCYSYGRLYSVKVNDCNVEICQSPVEYATVGSPIEL 133
Db 591 FA-----CICKEG-----WGGVTCAENLDDC-VGOCRN----- 617
Qy 134 AIGKKDQVDVITQOLPCEAEFVSDPE-----TPTSDG-----KLVWK-----IDRLG- 178
Db 618 ---GATCIDLVNDYRCACASGFTGRDCTDIDECATSPCRNGGECVDMVGKFCNICPLGY 674
Qy 179 AGDKC---KITVWVKPLKGGCFTA--ATVCAPELRSYTKCQO-PAIC----- 221
Db 675 SGLCEAEKENCPTSPCLEGHCLNTPGGYCHCPDRAGKHCEQLRPLCSQPPNEGCPA 734
Qy 222 -----IKQEGPDC-----ACLRCPV---CYKIEVVNTGSAIA 250
Db 735 NVSLATSATTTTTTTTATTTTRKMAKPSGLPCSGHSGCEMSDVGTFC-KCHVGHGTTCFE 793
Qy 251 RNVTVDNFVP-----DG-----YSHASQORVLS-----FNLGDMRPG-- 282
Db 794 HNLNECSPNCRNGGICLDGDDFTCECMGSGWTGKRCERATGCGYAGOCNGGTCMPGAP 853
Qy 283 DKV-----FTVEFCPQRRGQITNVATVTCGGHKCSANVTTVVNEPCVQVNSGA 333
Db 854 DKALOPHCRCAPGWTGLFCAE-----AIDOCROPCCHNGGT-----CE----SCA 894
Qy 334 DW-SYVCKPVEYSISVSNPGDLVLDVVIQDTPSGVTVLEAPGECICCNKVVVRKEMC 392
Db 895 GWFRVC-----AQGSGPDCRINNECSPPQCGGATCIGDIGYS-----IC 939
Qy 393 PGETLQKLVVKAQVGRFTNOVATVSESNGCTCTCAET-----TTHWKGAAATH 443
Db 940 P-----PGRHLRCEILLSDPKSACONASNTISPYTALNRSQNWLDIALTG 985

QY 444 MCVLDTNDPICGENTVYVYICVTNRGSAEDTNVSLILKFSKELQPIASSGPTKGTISGNT 503
Db 986 RTEDDENCACVCEN-----GTSRCTNLWCGLPNCYKVDPLSKSSNLGVCQKQHE 1035
QY 504 VVFDAL 509
Db 1036 VCVPAL 1041

RESULT 12

US-09-944-849-8
; Sequence 8, Application US/0944849
; GENERAL INFORMATION:
; APPLICANT: Nickoloff, Brian
; APPLICANT: Miele, Lucio
; TITLE OF INVENTION: METHOD AND REAGENTS FOR EPITHELIAL BARRIER FORMATION AND TREATMEN
; TITLE OF INVENTION: MALIGNANT AND BENIGN SKIN DISORDERS BY MODULATING THE NOTCH PATH
; FILE REFERENCE: 212583
; CURRENT APPLICATION NUMBER: US/09/944,849
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,614
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 1404
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-944-849-8

Query Match 4.4%; Score 130.5; DB 23; Length 1404;
Best Local Similarity 20.5%; Pred. No. 0.058;
Matches 124; Conservative 57; Mismatches 208; Indels 217; Gaps 32;

QY 14 LTSMASCFASGAGIEAAVAESLITKIVASAEKPAVPMTAKKRVLRNRKQVVEQKSRGA 73
Db 543 LTTTATAIIGSNLSSTALLAALTSVASTSLAIGPC-INAKECR-----NQP-----GS 590
QY 74 FCDKEFYPCCEGRCPVQAEQESCYGRLYSVKVNDCNVEICQSVPEATVGSVPYPIEL 133
Db 591 FA---CICKEG-----WGGVTCANLDDC-VQQRN----- 617
QY 134 AIGKDCVDVITQQLPCEAEFVSSDPE-----TTPTSDG-----KLWVK-----IDRLG- 178
Db 618 ---GATCIDLVNDYRCACASGFTGRDCEIDECATSPCRNGECVDMVGKFNCTCPUGY 674
QY 179 ASDKC---KITVWKPLKEGCGCFTA--ATVCACPELRSYTKCGQ-PAIC----- 221
Db 675 SGLCEBEAKENCTPSPCLGHCCLNTPEGYXCHCPDPDRAGKHCEQLRPLCSQPPCNEGCPA 734
QY 222 -----IKQGPDC-----ACLRCPV---CYKIEVWNTGSAIA 250
Db 735 NVSLATSATTTTTTTTATTTRKNAKPSGLPCSHGSGCEMSDVGTFC-KCHVGHTGTCPE 793
QY 251 RNYTVDNRPV-----DG-----YSHASGORVLS-----FNLGDMRPG-- 282
Db 794 HNLNECPNRCNGIGICLDGGDFTCECMGWTGKRCSRATGCTGACGQCGNGCTCMPAP 853
QY 283 DKKV-----FTVEFCPPORRGQITNATVTVCGGHKCSANVTTVNNEPCVQVNTSGA 333
Db 854 DKALQPHRCAPGWTGLFECAE-----AIDCRGQPCNHNGT-----CE-----SGA 894
QY 334 DW-SYCKPVEYSISVNPGLVLHDVVVITQPLPSGVTVLEAPGGICCNKVVWRIKMC 392
Db 895 GWFRCVC-----AQFGSPDCRINNVNCSQPQCGGATCIDGGVGYSC-----IC 939
QY 393 PGETLOFLVWKAQVPGRFTNQAVTSESNCICTSCAET-----TTHWKLATH 443
Db 940 P-----PGRHGLURCEILLSDPKSACONASNTISPYTALNRSQNWLDIALTG 985

QY 444 MCVLDTNDPICGENTVYVYICVTNRGSAEDTNVSLILKFSKELQPIASSGPTKGTISGNT 503

Db 986 RTEDDENCACVCEN-----GTSRCTNLWCGLPNCYKVDPLSKSSNLGVCQKQHE 1035
QY 504 VVFDAL 509
Db 1036 VCVPAL 1041

RESULT 13

PCT-US01-08631-46695
; Sequence 46695, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 46695
; LENGTH: 1483
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (75)..(121)
; OTHER INFORMATION: kv TRANSCRIPTASE REVERSE II ORF2 domain identified by
; OTHER INFORMATION: eMATRIX, accession number DM01354N, p-value=1.000e-40, raw sco
; OTHER INFORMATION: 13.17
; NAME/KEY: DOMAIN
; LOCATION: (783)..(1314)
; OTHER INFORMATION: Immunoglobulin domain identified by Pfam, accession name ig,
; OTHER INFORMATION: E-value=1.9e-38, Pfam score of 128.9
PCT-US01-08631-46695

Query Match 4.4%; Score 130.5; DB 1; Length 1483;
Best Local Similarity 19.3%; Pred. No. 0.063;
Matches 118; Conservative 90; Mismatches 217; Indels 185; Gaps 29;

QY 35 ITRKVASAETKPAV-----PWTAKKVR-----LVRRNKOPVEQKSRGACDKE--- 78
Db 626 VSQVVGCTTHPRDLRFNRIRIQGAFRRRLNLTLLNNQ-IKRIPSGAFEDLENLK 584
QY 79 FYPCGEGRCPVBAQESCYGRLYSVKVNDD---CNVEI---COSVPEYATVGS----- 126
Db 685 YLYLYKNEIQSIDRQAFKGLASLEQLRLDSNTLHCDCEILWLADLLKTYAESGNAQAAAI 744
QY 127 -PYPIETLAIKGRKDCVDVWVITQQLPCEAEFVSSDPEPTTPTSDGKLWVKIDRLGAGDKCKI 185
Db 745 CEYPRRI---QGRSVATITPEELNCERPRITSEPDADVTSGNTVYFTCRAGSNPKPEI 800
QY 186 TVWV--RPLKEGCGCFTAATVCACPELRSYTKGQPAICIKQEGPDCACLRCP-VCYKIEV 242
Db 801 -IWLNRKSLRHAADTGLSSRVDL-----TCLPNKG---SLLCPQTCGCPVK 846
QY 243 VNTGSAIARNVVDNVPDGYSHASQVRVLFNLGDMRPGDKKVFVTFPCPQRGGQITWV 302
Db 847 AKTALAAAVRTGSGTADHY-----MRPGEPLHSLSLTC---RLALPMI 888
QY 303 -----ATVT-----YCGGHKCSANVTTVNNEPCVQVNTSGADNSYVC---KPVEYSISV 349
Db 889 QSHLQATLTCLWLAVELTGTISVASNVTSLL-----FTCFCKIGTKVKLQVG 933
QY 350 NPGDL-----VLHD--VVIQDTLPSPGVTVLEA-----PGGEICCNKVVWRIKEMCPGET 396
Db 934 NELSNTKDSRLNLLDGGTLMQNTQETDQGIYQCMKNAVAGEVKTQEV-----T 982
QY 397 LQFKLVVKAQVPGRFT-NOVAVTSESNCICTSCAETTTTHWKLATHMVCVLTQNDP---- 452

QY 282 -----GDKKVTVECFORRGQITNATVTYCGHKCSA--NVTTVVNEPCVQVNISG 332
DB 413 IKHCHXGD-----CFF-CGQVCGLPNDTSKCHICKARCHEAVRVNRP-KEARPOA 462
QY 333 ADMSYVCKPVEYSISVSNEFDLVLDVVIQDTLPSTVTVLEAPGGEI-----CONKVVWRI 388
DB 463 KKEYKALP-----HPRCEEGVIVTCIGGHEVATWPCWNSKPTSC 502
QY 389 KEMCPGETLQFKL-VVKAQVPGFTNOAVTSESNCGT-CTSCAETTHMKGLAATHMCV 446
DB 503 QRSK---ARQLKCGNHKCSLVCHF---VPLPDMSAQTCANCEEGCTVPRPTGCIHACP 556
QY 447 LDTNDPICVGENTVYR 462
DB 557 KGCHPPPCAPCNFVIK 572

Search completed: May 25, 2002, 22:20:57
Job time: 509 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2002, 22:14:33 ; Search time 64.71 Seconds
(without alignments)
243.232 Million cell updates/sec

Title: US-09-523-647-2
Perfect score: 2951
Sequence: 1 MSKLRRVTVLALTSASC.....ILSSDTLTPVSDPENTHYV 556

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 116914 seqs, 28308587 residues

Total number of hits satisfying chosen parameters: 116914

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA-New.*

1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	120.5	4.1	716	5	US-09-312-283B-183	Sequence 183, App
2	118.5	4.0	1529	5	US-09-312-283B-396	Sequence 396, App
3	114.5	3.9	1036	6	US-10-121-049-142	Sequence 142, App
4	114.5	3.9	1036	6	US-10-121-050-142	Sequence 142, App
5	114.5	3.9	1036	6	US-10-121-053-142	Sequence 142, App
6	114.5	3.9	1036	6	US-10-121-043-142	Sequence 142, App
7	114.5	3.9	1036	6	US-10-121-044-142	Sequence 142, App
8	114.5	3.9	1036	6	US-10-121-047-142	Sequence 142, App
9	114.5	3.9	1036	6	US-10-121-054-142	Sequence 142, App
10	114.5	3.9	1036	6	US-10-121-056-142	Sequence 142, App
11	114.5	3.9	1036	6	US-10-121-057-142	Sequence 142, App
12	114.5	3.9	1036	6	US-10-121-058-142	Sequence 142, App
13	114.5	3.9	1036	6	US-10-121-060-142	Sequence 142, App
14	114.5	3.9	1036	6	US-10-121-063-142	Sequence 142, App
15	114.5	3.9	1036	6	US-10-123-108-142	Sequence 142, App
16	114.5	3.9	1036	6	US-10-123-154-142	Sequence 142, App
17	114.5	3.9	1036	6	US-10-123-156-142	Sequence 142, App
18	114.5	3.9	1036	6	US-10-123-157-142	Sequence 142, App
19	114.5	3.9	1036	6	US-10-123-212-142	Sequence 142, App
20	114.5	3.9	1036	6	US-10-123-213-142	Sequence 142, App
21	114.5	3.9	1036	6	US-10-123-109-142	Sequence 142, App
22	114.5	3.9	1036	6	US-10-121-041-142	Sequence 142, App
23	114.5	3.9	1036	6	US-10-121-045-142	Sequence 142, App
24	114.5	3.9	1036	6	US-10-121-046-142	Sequence 142, App
25	114.5	3.9	1036	6	US-10-121-031-142	Sequence 142, App
26	114.5	3.9	1036	6	US-10-121-040-142	Sequence 142, App

ALIGNMENTS

RESULT 1

US-09-312-283B-183

; Sequence 183, Application US/09312283B

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.

; APPLICANT: Strachan, Lorna

; APPLICANT: Sleeman, Matthew

; APPLICANT: Onrust, Rene

; APPLICANT: Mullison, James G.

; APPLICANT: Kumble, Krishanend D. Isolated from Skin Cells

; TITLE OF INVENTION: Compositions and Methods for their Use

; FILE REFERENCE: 11000.1011c2

; CURRENT APPLICATION NUMBER: US/09/312.283B

; CURRENT FILING DATE: 1999-05-14

; NUMBER OF SEQ ID NOS: 425

; SOFTWARE: FASTSEQ for Windows Version 4.0

; SEQ ID NO 183

; LENGTH: 716

; TYPE: PRT

; ORGANISM: Mouse

US-09-312-283B-183

Query Match 4.1%; Score 120.5; DB 5; Length 716;

Best Local Similarity 20.4%; Pred. No. 0.17; 126; Indels 191; Gaps 24;

Matches 92; Conservative 42; Mismatches 191; Gaps 24;

Qy 72 GAFCDKEFYPCERGCQ-----EVEAOQESC---YGRLY-----SVKVNDD 109

Db 313 GEHCIDFDQDNKCKNGAHCCTDAVNGYTCVCPGYSGLFCFSPPMVFLRTSPCDNFD 372

Qy 110 C-----NVEICQSVPEYATVCSVPYTEILAIGKKDC-----VDVWITQQLPCEAE 154

Db 373 CONGAOCIIIRVNEPICQCLPGY-----LGEK-CEKLVSVSILVNKESYLQIP 418

Qy 155 FVSSDPEF-----TPTSDGKLVKIDELGAGDKCKITVWVKPLKGGCCFATVACAP 207

Db 419 SAKVRPQNTITLQIATDESGILLYK-----GDKDHAV--ESIEGI-----458

Qy 208 ELRSYTKCGQFATCIKQEGPDCACLCRCPCVKYKIEVNTGS-----AIARVNTVDNP 258

Db 459 -RASYPDTGSHFASAI-----YSVETINDGNFHFIVELLTLOSLSLSDGG 502

Qy 259 VPQGYSHASQRYLSEFN---LGDMPGDKKVFTEFCPQRGQITNVATVTCGGHKCS 314

Db 503 SPKIITNLKSKQTLNFDSPLYVGGM-PGKNNVASLRQAPGQNG-----TSFHCCI 551

Qy 315 ANVTTVNEPCVQVNVISGADWSVCKPVEYSISVSNPGDLVLHDVVIODTLPSGVTVLEA 374

27 114.5 3.9 1036 6 US-10-121-048-142 Sequence 142, App

28 114.5 3.9 1036 6 US-10-121-052-142 Sequence 142, App

29 114.5 3.9 1036 6 US-10-121-061-142 Sequence 142, App

30 114.5 3.9 1036 6 US-10-121-042-142 Sequence 142, App

31 114.5 3.9 1036 6 US-10-121-055-142 Sequence 142, App

32 114.5 3.9 1036 6 US-10-121-059-142 Sequence 142, App

33 114.5 3.9 1036 6 US-10-124-822-142 Sequence 142, App

34 114.5 3.9 1036 6 US-10-123-903-142 Sequence 142, App

35 114.5 3.9 1036 6 US-10-124-817-142 Sequence 142, App

36 114.5 3.9 1036 6 US-10-124-819-142 Sequence 142, App

37 114.5 3.9 1036 6 US-10-124-823-142 Sequence 142, App

38 114.5 3.9 1036 6 US-10-125-704-142 Sequence 142, App

39 114.5 3.9 1036 6 US-10-123-215-142 Sequence 142, App

40 114.5 3.9 1036 6 US-10-123-235-142 Sequence 142, App

41 114.5 3.9 1036 6 US-10-123-236-142 Sequence 142, App

42 114.5 3.9 1036 6 US-10-123-261-142 Sequence 142, App

43 114.5 3.9 1036 6 US-10-123-322-142 Sequence 142, App

44 114.5 3.9 1036 6 US-10-124-820-142 Sequence 142, App

45 114.5 3.9 1036 6 US-10-123-771-142 Sequence 142, App

Db 552 RN-----LVINSELODFKRV--PMOTGI-----L 573
Qy 375 PGGEICCNKVVWRIKEMCPGETLOFKLVVKAQVPGRTNOVAVTSBNGCTCTSCAETTT 434
Db 574 PGGEPCHKV-----CAHGTGQ-----PSSQSGFTCECE--- 603
Qy 435 HWKGLAATHMCVLDNDPTCVGENTVYRICV 465
Db 604 GWMG----PLCDORTNDP--CLGNKCVHGTCL 629
RESULT 2
US-09-312-283B-396
; Sequence 396, Application US/09312283B
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312.283B
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 396
; LENGTH: 1529
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283B-396

Query Match 4.0%; Score 118.5; DB 5; Length 1529;
Best Local Similarity 20.3%; Pred. No. 0.6;
Matches 94; Conservative 39; Mismatches 115; Indels 215; Gaps 25;
Qy 72 GAFCDKEFYPCBGRQ-----PVEAQOESC---YGRLY-----SVKVNDD 109
Db 1070 GEHCDIDFDCQDNKNGAHCTDAVNGYTCVCEGYSGLFCEFPMPVLLRTSPCDNFD 1129
Qy 110 C-----NVEICQSVPEYATVGSPPYPIEILAIKDKCDVDVITQQLPCEAEFVSSD 159
Db 1130 CONGAQCIIRVNEPICQLPGY-----LGEK-CEKLV-----SVNEVNE 1168
Qy 160 -----PFT-----TPTSDGKLVWKIDRLGAGDKCKITVWVKPLKESCCFTAA 201
Db 1169 SYLQIPSAKYRPTNITLQIATDEDSGLLYK-----GKDHIAV----- 1208
Qy 202 TVCACPEL-----RSYTKGGOPATCIKQEGPDCACLRCPVCYKIEVVNTGS----- 247
Db 1209 -----ELYGRVRSYDTGSHPSAI-----YSVETINDGNFHVLELT 1247
Qy 248 -AIARNYVDNVPDGYSHASQORVLSFN-----LGDMPRGDKVFTVEFCFQRRGQITNV 302
Db 1248 LDSSLSLSVDGGSPKIIITNLRSKOSTLNFDSPLYVGGM--PGKNNVASLRQAPGONG----- 1301
Qy 303 ATVTYCGCHKCSANVTTVVNEPCVOVNISGADWSYVCKPVEYSISVSNPGDLVLDHWIQ 362
Db 1302 -----TSFHCICRN-----LYINSELODFKRV--PMOTGI----- 1329
Qy 363 DTLPSGVTVLEAPGGEICCNKVVWRIKEMCPGETLOFKLVVKAQVPGRTNOVAVTSBNG 422
Db 1330 -----LFGCEPCHKV-----CAHGTGQ-----PSSQ 1351
Qy 423 CGTCTCAETTTWKGLAATHMCVLDNDPTCVGENTVYRICV 465
Db 1352 SGFTCECE---GWMG-----PLCDORTNDP--CLGNKCVHGTCL 1386

RESULT 3
US-10-121-049-142
; Sequence 142, Application US/10121049
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C17
; CURRENT APPLICATION NUMBER: US/10/121.049
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 142
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-049-142

Query Match 3.9%; Score 114.5; DB 6; Length 1036;
Best Local Similarity 19.2%; Pred. No. 0.75;
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;
Qy 75 CDKEFYPCBGRQVPEAQOESCYGRLYSVK-----VNDCHVETCQ 116
Db 376 CERYVVP--EGECCPV-----CEDPVYPNPNFAGCYANGLLHAHDRWREDDCT--FCQ 425
Qy 117 SVPEVATGSPYPIEILAIKDKCDVDVITQ-----OLPCEAEFVSSDP---ETTPSTD 167
Db 426 CVN-----GERHCVATVCGTCTNPKVPGCECPVCEPTIITVDPPAC 469
Qy 168 GKLV-----WKIDRLGAGDKCKITVWV-----KPLKEG---C-----FTAAT 202
Db 470 GELSNTLTGKDCINGFKRDHG---CRTCCOCLNTEELCSEKQCGCTLNCPPGFLTDAQ 525
Qy 203 VCACPELRSYTKGGOPATCIK-----QEGPD-CACLRCP--VCYKIEVVNTGSAIA 250
Db 526 NCIECRRPRPKRPIICDKYCPGLGLKNKHGCDICRCKKCPCLSCSI----- 575
Qy 251 RNVVDNVPDGYSHASQORVLSFNGLDMRPGDKVFTVEFCFQRRGQITNVATVYTCGG 310
Db 576 -----CPLGQDQSHGCLICKCRE 594
Qy 311 HKCSANVTTVNPEPCVQV-----NISGADWSYVCKPV-----EYSISVSNP---GDLVL 356
Db 595 ASASAG--PPLISGTCLTVDGHGHHKNEESWHDGCRCYCLNGREMCALITCPVPACGNPTI 653
Qy 357 H-----DVVLODTLPSGVTVLEAPGGEICCNKVVWRI-----KE 390
Db 654 HPGCCPSCADDFVQKRPSTPSICHAPGGEYFVEGETHWNIDSTQCTCHSGRVLCE 713
Qy 391 MCPGETLOFKLVVKAQVPGRTNOVAVTSBNGCTCTSCAETTTWKGLAATHMCVLDTN 450
Db 714 VCP-----PLLCNPSR-----TQDSCCPQCTDQFPRPSLRNNSVNPYCNKDEG 758
Qy 451 DPTICGENTVYRICV 466

```
Db 759 DIFLAAESWKPDVCTS 774

RESULT 4
US-10-121-050-142
: Sequence 142, Application US/10121050
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: DeForge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P33301C20
: CURRENT APPLICATION NUMBER: US/10/121,050
: CURRENT FILING DATE: 2002-04-12
: Prior Application removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 142
: LENGTH: 1036
: TYPE: PRT
: ORGANISM: Homo Sapien
US-10-121-050-142

Query Match 3.9%; Score 114.5; DB 6; Length 1036;
Best Local Similarity 19.2%; Pred. No. 0.75;
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;

Qy 75 CDKEFYPCERGRCQPEAQOESCGRLYSVK-----VNDCCNVEICQ 116
Db 376 CERYVP--EGECCPV-----CEDPVYFNNPAGCYANGLILAHGDRWRDDCT--FCQ 425
Qy 117 SVPEYATVGSPPYPIELTAIGKKDCVDVVITO-----OLPCEAEFYSSDP---ETPTSD 167
Db 426 CVN-----GERHCYATVCGQCTNPNVKVFGECPCVEEPTIITVDPPAC 469
Qy 168 GKL-----WKIDRLGAGDKCKITVWV-----KPLKEGC---C-----FTAA 202
Db 470 GELSNCTLTGKDCINGFKRDHNG-----CRTCCINTELCSERKQGTCLNCPFGFLTDAQ 525
Qy 203 VCACPELRSYTKGQPAICIK-----QEGPD-CACLRCP--VCYKIEVNTGSAIA 250
Db 526 NCEICECRPRPKCRPIICDKYCPGLGLLNKHHGCDICRCKKCPELSCKSI-----575
Qy 251 RNVTVDNPNPDGYSHASGQVLSFNLGDMRPGDKKVFTEFCPQRRGQITNATVTCGG 310
Db 576 -----CPLGFOODSHGCLICKCRE 594
Qy 311 HKCSANVTTVNPEPCVOV-----NISGADNSYVCKPV-----EYSISVSNP-----GDLVL 356
Db 595 ASASAG--PPILSGTCLTVGHHKNEESHDGCRECYCLNGREMCALITCPVACGNPTI 653
Qy 357 H-----DVVIQDTLPISGVTVLEAPGGEICCNKVVWRI-----KE 390
Db 654 HPGQCCPSCADDFWQKPELSTPSICHAPGGEYFVEGETWNIDSCOTCHSGRVLCE 713
Qy 301 MCPGETLQFKLVKVAQVGRFTNOVAVTSNCGTCTSCAETTHWKGAAATHMCVLDTN 450
Db 714 VCP-----PILCQNPSSR-----TQDSCCPQCTDQPFPSLRNNSVFNPKNDEG 758
```

```
Db 759 DIFLAAESWKPDVCTS 774

RESULT 5
US-10-121-053-142
: Sequence 142, Application US/10121053
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: DeForge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P33301C23
: CURRENT APPLICATION NUMBER: US/10/121,053
: CURRENT FILING DATE: 2002-04-12
: Prior Application removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 142
: LENGTH: 1036
: TYPE: PRT
: ORGANISM: Homo Sapien
US-10-121-053-142

Query Match 3.9%; Score 114.5; DB 6; Length 1036;
Best Local Similarity 19.2%; Pred. No. 0.75;
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;

Qy 75 CDKEFYPCERGRCQPEAQOESCGRLYSVK-----VNDCCNVEICQ 116
Db 376 CERYVP--EGECCPV-----CEDPVYFNNPAGCYANGLILAHGDRWRDDCT--FCQ 425
Qy 117 SVPEYATVGSPPYPIELTAIGKKDCVDVVITO-----OLPCEAEFYSSDP---ETPTSD 167
Db 426 CVN-----GERHCYATVCGQCTNPNVKVFGECPCVEEPTIITVDPPAC 469
Qy 168 GKL-----WKIDRLGAGDKCKITVWV-----KPLKEGC---C-----FTAA 202
Db 470 GELSNCTLTGKDCINGFKRDHNG-----CRTCCINTELCSERKQGTCLNCPFGFLTDAQ 525
Qy 203 VCACPELRSYTKGQPAICIK-----QEGPD-CACLRCP--VCYKIEVNTGSAIA 250
Db 526 NCEICECRPRPKCRPIICDKYCPGLGLLNKHHGCDICRCKKCPELSCKSI-----575
Qy 251 RNVTVDNPNPDGYSHASGQVLSFNLGDMRPGDKKVFTEFCPQRRGQITNATVTCGG 310
Db 576 -----CPLGFOODSHGCLICKCRE 594
Qy 311 HKCSANVTTVNPEPCVOV-----NISGADNSYVCKPV-----EYSISVSNP-----GDLVL 356
Db 595 ASASAG--PPILSGTCLTVGHHKNEESHDGCRECYCLNGREMCALITCPVACGNPTI 653
Qy 357 H-----DVVIQDTLPISGVTVLEAPGGEICCNKVVWRI-----KE 390
Db 654 HPGQCCPSCADDFWQKPELSTPSICHAPGGEYFVEGETWNIDSCOTCHSGRVLCE 713
Qy 301 MCPGETLQFKLVKVAQVGRFTNOVAVTSNCGTCTSCAETTHWKGAAATHMCVLDTN 450
Db 714 VCP-----PILCQNPSSR-----TQDSCCPQCTDQPFPSLRNNSVFNPKNDEG 758
```

```
QY 391 MCPGETLQFLVKAQVPGFRFTNOAVTSESNGTCTSCAETTHWKGLAATHMCVLDTN 450
Db 714 VCP-----PILCONPSR-----TODSCCPQCTDQPPRPSLSRNNSPVNYCKNDEG 758
QY 451 DPICVGTNTVYRICVT 466
Db 759 DIFLAESWKPDPVCTS 774

RESULT 6
US-10-121-043-142
; Sequence 142, Application US/10121043
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C15
; CURRENT APPLICATION NUMBER: US/10/121,043
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 142
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-043-142
```

Query Match 3.9%; Score 114.5; DB 6; Length 1036;
Best Local Similarity 19.2%; Pred. No. 0.75;
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;

```
QY 75 CDKEFYPCGEGRCOPVEAQOQESCYGRLYSVK-----VNDCCNVEICQ 116
Db 376 CERYVVP--EGECCPV-----CEDPVYFPNPNAGYANGLIHAHGDWRREDDCT--FCQ 425
QY 117 SVPEYATVGSYPYPIEILAIKGDVVDWITQ-----OLPCEAEFVSSDP---ETPTSD 167
Db 426 CVN-----GERHCVATVCGTCTNPVKVPGECPCVEEPTIITVDPPAC 459
QY 168 GKLV-----WKIDRLGAGDKCKITVWV-----KPLKEGC---C-----FTAAT 202
Db 470 GELSNCILTGKDCINGFKRDHNG-----CRTCQCINTEELCSERKQCTLNCPEGLFLDAQ 525
QY 203 VCACPELSYTKCQQAICIK-----QEGPD-CACLRCP--VCYKIEVNTGSAIA 250
Db 526 NCEICECRPRPKCRPIICDKYCPGLGLLNKKGCDICRCKCPELSCSKI-----575
QY 251 RNVTVDPNPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPQRRGQITNVATVYCGG 310
Db 576 -----CPLGFQDSSHGCLICKRE 594
QY 311 HKCSANVTTVNNEPCVQV-----NISGADWSYVCKPV-----EYSISVSNP---GDLVL 356
Db 595 ASASAG-PPILSGTCLTVDGHGHHKNESSHDCRECYCLNGREMCALITCPVACGNPTI 653
QY 357 H-----DVIQDTLPVSGTVLEAPGGEICCNKVVWRI-----KE 390
```

```
Db 654 HPQCCPSCADDFVQKPELSTPSCIHAPGGEYFVEGETWNIDSCQTQCTCHSRVUCETE 713
QY 391 MCPGETLQFLVKAQVPGFRFTNOAVTSESNGTCTSCAETTHWKGLAATHMCVLDTN 450
Db 714 VCP-----PILCONPSR-----TODSCCPQCTDQPPRPSLSRNNSPVNYCKNDEG 758
QY 451 DPICVGTNTVYRICVT 466
Db 759 DIFLAESWKPDPVCTS 774

RESULT 7
US-10-121-044-142
; Sequence 142, Application US/10121044
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C15
; CURRENT APPLICATION NUMBER: US/10/121,044
; CURRENT FILING DATE: 2002-04-11
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 142
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-044-142
```

Query Match 3.9%; Score 114.5; DB 6; Length 1036;
Best Local Similarity 19.2%; Pred. No. 0.75;
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;

```
QY 75 CDKEFYPCGEGRCOPVEAQOQESCYGRLYSVK-----VNDCCNVEICQ 116
Db 376 CERYVVP--EGECCPV-----CEDPVYFPNPNAGYANGLIHAHGDWRREDDCT--FCQ 425
QY 117 SVPEYATVGSYPYPIEILAIKGDVVDWITQ-----OLPCEAEFVSSDP---ETPTSD 167
Db 426 CVN-----GERHCVATVCGTCTNPVKVPGECPCVEEPTIITVDPPAC 459
QY 168 GKLV-----WKIDRLGAGDKCKITVWV-----KPLKEGC---C-----FTAAT 202
Db 470 GELSNCILTGKDCINGFKRDHNG-----CRTCQCINTEELCSERKQCTLNCPEGLFLDAQ 525
QY 203 VCACPELSYTKCQQAICIK-----QEGPD-CACLRCP--VCYKIEVNTGSAIA 250
Db 526 NCEICECRPRPKCRPIICDKYCPGLGLLNKKGCDICRCKCPELSCSKI-----575
QY 251 RNVTVDPNPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCPQRRGQITNVATVYCGG 310
Db 576 -----CPLGFQDSSHGCLICKRE 594
QY 311 HKCSANVTTVNNEPCVQV-----NISGADWSYVCKPV-----EYSISVSNP---GDLVL 356
```


Db 595 ASASAG-PPILSGTCLTVGHHHKNESWHDGCRACYCLNGREMCALITCPVPACGNPTI 653
QY 357 H-----DVIQDTLPSSQVTVLEAPGGEICCNKVVWRI-----KE 390
Db 654 HPGCCPSCADDFVQKPELSPSICHAPGGEYFVEGETWNIDSCQTCTCHSGRVLCE 713
QY 391 MCPGETLQFLVKAQVPGRETNOVAVTSBNGCTCTSCAETTHWKGLAATHMVCVLDTN 450
Db 714 VCP-----PLLCNPSR-----TQDSCCPQCTDQFPRLSRNNSVNYCKNDEG 758
QY 451 DPICVGVNTVYRICVT 466
Db 759 DIFLAESWKPDOVCTS 774

RESULT 8
US-10-121-047-142
; Sequence 142, Application US/10121047
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C4
; CURRENT APPLICATION NUMBER: US/10/121.047
; CURRENT FILING DATE: 2002-04-11
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 142
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-047-142

Query Match 3.9%; Score 114.5; DB 6; Length 1036;
Best Local Similarity 19.2%; Pred. No. 0.75;
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;
QY 75 CDKEFYPCBGRCPVFAQGESCYGRLYSVK-----VNDGCVETCQ 116
Db 376 CERYVP--EGECCPV-----CEDPVYFNNPAGCYANGLIHAGDRWREDCT--FCQ 425
QY 117 SVPEYATVGSPPYIEILAIKKDCVDVITQ-----OLPCEAEFVSSDP---ETPTSD 167
Db 426 CVN-----GERHCVATVCGQTCTNPVKVGECCPCEPEPTIITVDPFAC 469
QY 168 GKLV-----WKIDRLGAGDKCKITVWV-----KPLKEGC---C---FTAA 202
Db 470 GELSNCTLTGRDCINGFKRDHNG---CRTQCINTBELCSERKQGTCLNCPFGELTDAQ 525
QY 203 VCACPELRSYTKCGQPAICIK-----OEGPD-CACLRCP--VCYKIEVVNTGSAIA 250
Db 526 NCEICECRPRKCRPIIDCKYCPGLGLLKNKHGDCICRCKCPELSCSKI----- 575
QY 251 RNVTVDNPVDSYSHASQGVRLSPNLGDMRPDKKVFTEVBCPQRRGQITNATVTCGG 310
Db 576 -----CPLGFOQDSHGCLICKRE 594

QY 311 HKCSANVTTVWNEPCVQV-----NISGADWSYCKPV-----EYSISVSNP-----GDLVL 356
Db 595 ASASAG-PPILSGTCLTVGHHHKNESWHDGCRACYCLNGREMCALITCPVPACGNPTI 653
QY 357 H-----DVIQDTLPSSQVTVLEAPGGEICCNKVVWRI-----KE 390
Db 654 HPGCCPSCADDFVQKPELSPSICHAPGGEYFVEGETWNIDSCQTCTCHSGRVLCE 713
QY 391 MCPGETLQFLVKAQVPGRETNOVAVTSBNGCTCTSCAETTHWKGLAATHMVCVLDTN 450
Db 714 VCP-----PLLCNPSR-----TQDSCCPQCTDQFPRLSRNNSVNYCKNDEG 758
QY 451 DPICVGVNTVYRICVT 466
Db 759 DIFLAESWKPDOVCTS 774

RESULT 9
US-10-121-054-142
; Sequence 142, Application US/10121054
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C4
; CURRENT APPLICATION NUMBER: US/10/121.054
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 142
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-054-142

Query Match 3.9%; Score 114.5; DB 6; Length 1036;
Best Local Similarity 19.2%; Pred. No. 0.75;
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;
QY 75 CDKEFYPCBGRCPVFAQGESCYGRLYSVK-----VNDGCVETCQ 116
Db 376 CERYVP--EGECCPV-----CEDPVYFNNPAGCYANGLIHAGDRWREDCT--FCQ 425
QY 117 SVPEYATVGSPPYIEILAIKKDCVDVITQ-----OLPCEAEFVSSDP---ETPTSD 167
Db 426 CVN-----GERHCVATVCGQTCTNPVKVGECCPCEPEPTIITVDPFAC 469
QY 168 GKLV-----WKIDRLGAGDKCKITVWV-----KPLKEGC---C---FTAA 202
Db 470 GELSNCTLTGRDCINGFKRDHNG---CRTQCINTBELCSERKQGTCLNCPFGELTDAQ 525
QY 203 VCACPELRSYTKCGQPAICIK-----OEGPD-CACLRCP--VCYKIEVVNTGSAIA 250
Db 526 NCEICECRPRKCRPIIDCKYCPGLGLLKNKHGDCICRCKCPELSCSKI----- 575

```

Qy 251 RNVTDNPVPGYSHASQORVLSFNLGDMRPGDKVTFVEFCPPORRGOITNATVTVYCGG 310
Db 576 -----CPLGFQODSHGCLICKRE 594

Qy 311 HKCSANVTTVVNEPCVOV-----NISGADMSYVCKPV-----EYSISVSNP-----GDLVL 356
Db 595 ASASAG-PPILSGTCLTVDGHHHKNESWHDGCRVCYCLNGREMCALITCPVPACGNPTI 653
Qy 357 H-----DVIQDTLPVSGVTVLEAPGGEICCNKVVWRI-----EYSISVSNP-----GDLVL 356
Db 654 HPQCPCSCADDFVQKPELSTPSICHAPGGEYFVEGETWNIDSTQCTCHSGRVLCE 713
Qy 391 MCPGETLQFLVYKAQVGRFTNOVAVTSESNGCTCTSCAETTHWKGLAATHMVCVLDTN 450
Db 714 VCP-----PLLCNFSR-----TQDSCCPQCTDQFPFRLSRNNSVNPYCKNDEG 758

Qy 451 DPICVGVNTVYRICVT 466
Db 759 DIFLAAESWKPDVCTS 774

```

RESULT 10

```

US-10-121-056-142
: Sequence 142, Application US/10121056
: GENERAL INFORMATION:

```

```

: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: DeForge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3330RIC23
: CURRENT APPLICATION NUMBER: US/10/121,056
: CURRENT FILING DATE: 2002-04-12
: Prior Application removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 142
: LENGTH: 1036
: TYPE: PRT
: ORGANISM: Homo Sapien
US-10-121-056-142

```

```

Query Match 3.9%; Score 114.5; DB 6; Length 1036;
Best Local Similarity 19.2%; Pred. No. 0.75;
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;

```

```

Qy 75 CDKEFYPCBEGRCQPVFAQOQESCYGRLYSVK-----VNDDCNVEICQ 116
Db 376 CERYIYP--EGECCPV-----CEDPYFPNNPAGCYANGLILANGDRWREDDCT--FCQ 425
Qy 117 SVPEYATVSGPYPIELAIKKKDCVDVITQ-----QLPCAERFVSSDP---ETTPSTD 167
Db 426 CVN-----GERHCVATVCGQCTNPNKVPKGCCPCVCEPTIITVDPPAC 469
Qy 168 GKLV-----WKIDRLGAGDKCKITVWV-----KPLKECC---C-----FTAA 202
Db 470 GELSNTLTGRDCINGEVRDRNG-----CRTCOCINTEELCSRKQKQCTLNCFPGLTDAQ 525
Qy 203 VCACPELRSYTKCGOPAIKIC-----QEGPD-CACLRCP--VCYKIEWVNTGSAIA 250

```

```

Db 526 NCEICECRPRPKCRPIICDKYCPILGLLKNKHGCDICRCKKPELSCSKI----- 575
Qy 251 RNVTDNPVPGYSHASQORVLSFNLGDMRPGDKVTFVEFCPPORRGOITNATVTVYCGG 310
Db 576 -----CPLGFQODSHGCLICKRE 594
Qy 311 HKCSANVTTVVNEPCVOV-----NISGADMSYVCKPV-----EYSISVSNP-----GDLVL 356
Db 595 ASASAG-PPILSGTCLTVDGHHHKNESWHDGCRVCYCLNGREMCALITCPVPACGNPTI 653
Qy 357 H-----DVIQDTLPVSGVTVLEAPGGEICCNKVVWRI-----EYSISVSNP-----GDLVL 356
Db 654 HPQCPCSCADDFVQKPELSTPSICHAPGGEYFVEGETWNIDSTQCTCHSGRVLCE 713
Qy 391 MCPGETLQFLVYKAQVGRFTNOVAVTSESNGCTCTSCAETTHWKGLAATHMVCVLDTN 450
Db 714 VCP-----PLLCNFSR-----TQDSCCPQCTDQFPFRLSRNNSVNPYCKNDEG 758

Qy 451 DPICVGVNTVYRICVT 466
Db 759 DIFLAAESWKPDVCTS 774

```

RESULT 11

```

US-10-121-057-142
: Sequence 142, Application US/10121057
: GENERAL INFORMATION:

```

```

: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: DeForge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3330RIC13
: CURRENT APPLICATION NUMBER: US/10/121,057
: CURRENT FILING DATE: 2002-04-12
: Prior Application removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 142
: LENGTH: 1036
: TYPE: PRT
: ORGANISM: Homo Sapien
US-10-121-057-142

```

```

Query Match 3.9%; Score 114.5; DB 6; Length 1036;
Best Local Similarity 19.2%; Pred. No. 0.75;
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;

```

```

Qy 75 CDKEFYPCBEGRCQPVFAQOQESCYGRLYSVK-----VNDDCNVEICQ 116
Db 376 CERYIYP--EGECCPV-----CEDPYFPNNPAGCYANGLILANGDRWREDDCT--FCQ 425
Qy 117 SVPEYATVSGPYPIELAIKKKDCVDVITQ-----QLPCAERFVSSDP---ETTPSTD 167
Db 426 CVN-----GERHCVATVCGQCTNPNKVPKGCCPCVCEPTIITVDPPAC 469
Qy 168 GKLV-----WKIDRLGAGDKCKITVWV-----KPLKECC---C-----FTAA 202

```

Db 470 GELSNCTLTGKDCINGFKRDHNG-----CRTCQCINTFELCSEKQGGCTLNCPPGFLITDAQ 525
Qy 203 VCACPELRSYTKGQPAICIK-----OEGPD-CACLRCP--VCYKIEVNTGSAIA 250
Db 526 NCEICECRPRPKRPIICDKYCPLGLLKNKHGCDICRCKKPELSCSKI-----575
Qy 251 RNVTVDPVDPGYSHASGORVLSFNLGDMRPGDKKVFTEFCPQRRGQITNVATVYCGG 310
Db 576 -----CPLGFQODSHGCLICKCRE 594
Qy 311 HKCSANVTTVNPPCVQV-----NISGADWSYVCKPV-----EYSISVSNP-----GDLVL 356
Db 595 ASASAG-PPILSGTCLTVDGHHKNEESWHDGCRECYCLNGREMCALITCPVPACGNPTI 653
Qy 357 H-----DVIODTLPSGTVVLEAPGGEICCNKVVWRI-----KE 390
Db 654 HPGQCCPSCADDFVQKPELSPSICHAPGGEYFVEGETWNIIDSCOTCTCHSGRVLCE 713
Qy 391 MCPGETLOFLVYKAQVGRFTNQAVTSENCGTCTSCAETTHWKGLAATHMVCVLDN 450
Db 714 VCP-----PLLCQNPSR-----TQDSCCPQCTDQPPRPSLSRNNSVPNYCKNDEG 758
Qy 451 DPICVGVNTVYRICVT 466
Db 759 DIFLAESWKPDCVCTS 774
RESULT 12
US-10-121-058-142
Sequence 142, Application US/10121058
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Flivaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C12
CURRENT APPLICATION NUMBER: US/10/121,058
CURRENT FILING DATE: 2002-04-12
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 142
LENGTH: 1036
TYPE: PRT
ORGANISM: Homo Sapien
US-10-121-058-142

Query Match 3.9%; Score 114.5; DB 6; Length 1036;
Best Local Similarity 19.2%; Pred. No. 0.75;
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;
Qy 75 CKEFEYPCGEGRCPVFAOQESCYRLYSVK-----VNDCCNVEICQ 116
Db 376 CERYVFP--EGECCPV-----CEDPVYFPNNPAGCYANGLILAHGRWRREDCT--FCQ 425
Qy 117 SVPEYATVGSPIYETILAIKDKCDVDWITO-----QLPCAEEFVSDP---ETPTSD 167
Db 426 CVN-----GERHCVATVCGQTCTNPVKVPGCECPVCEEPTITVDPPAC 469

Qy 168 GKLY-----WKIDRLGAGDKCKITVWV-----KPLKEGC-----C-----FTAAT 202
Db 470 GELSNCTLTGKDCINGFKRDHNG-----CRTCQCINTFELCSEKQGGCTLNCPPGFLITDAQ 525
Qy 203 VCACPELRSYTKGQPAICIK-----OEGPD-CACLRCP--VCYKIEVNTGSAIA 250
Db 526 NCEICECRPRPKRPIICDKYCPLGLLKNKHGCDICRCKKPELSCSKI-----575
Qy 251 RNVTVDPVDPGYSHASGORVLSFNLGDMRPGDKKVFTEFCPQRRGQITNVATVYCGG 310
Db 576 -----CPLGFQODSHGCLICKCRE 594
Qy 311 HKCSANVTTVNPPCVQV-----NISGADWSYVCKPV-----EYSISVSNP-----GDLVL 356
Db 595 ASASAG-PPILSGTCLTVDGHHKNEESWHDGCRECYCLNGREMCALITCPVPACGNPTI 653
Qy 357 H-----DVIODTLPSGTVVLEAPGGEICCNKVVWRI-----KE 390
Db 654 HPGQCCPSCADDFVQKPELSPSICHAPGGEYFVEGETWNIIDSCOTCTCHSGRVLCE 713
Qy 391 MCPGETLOFLVYKAQVGRFTNQAVTSENCGTCTSCAETTHWKGLAATHMVCVLDN 450
Db 714 VCP-----PLLCQNPSR-----TQDSCCPQCTDQPPRPSLSRNNSVPNYCKNDEG 758
Qy 451 DPICVGVNTVYRICVT 466
Db 759 DIFLAESWKPDCVCTS 774
RESULT 13
US-10-121-060-142
Sequence 142, Application US/10121060
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Flivaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C21
CURRENT APPLICATION NUMBER: US/10/121,060
CURRENT FILING DATE: 2002-04-12
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 142
LENGTH: 1036
TYPE: PRT
ORGANISM: Homo Sapien
US-10-121-060-142

Query Match 3.9%; Score 114.5; DB 6; Length 1036;
Best Local Similarity 19.2%; Pred. No. 0.75;
Matches 95; Conservative 45; Mismatches 155; Indels 201; Gaps 24;
Qy 75 CKEFEYPCGEGRCPVFAOQESCYRLYSVK-----VNDCCNVEICQ 116
Db 376 CERYVFP--EGECCPV-----CEDPVYFPNNPAGCYANGLILAHGRWRREDCT--FCQ 425

Qy	117	SYPEATVGSVPPIELIALGKKCCDVVITQ-----QLPCEAEVSSDP-----ETTPSTD	167
Db	426	CVN-----GERHCVAIVCGQTCTNPVKVPGECCPVCBEPTIITVDPAC	469
Qy	168	GLIV-----WKIDRLGAGDKCKITVW-----KPLKGC---C-----FTAAT	202
Db	470	GELSNCITLTKDCINGFKRHNG-----CRTCCOCTEELSCSERKGCTLCPFGFLTDAQ	525
Qy	203	VCACPCLSYNKKGCPATCIK-----OEGPD-CACLSCP---VCYKIEVNTGSAIA	250
Db	526	NCEICECPRPKKRPIICDKYCPGLLLKNKHGCDICRCKKPELSCSKI-----	575
Qy	251	RNVTVNVPDGYSHASQORVLSFNLGDMRPGKVTVVEFCPPRRGOITNVAIVTCGG	310
Db	576	-----CPGFQDSHGCLICKRE	594
Qy	311	HKCSANVTTVNPEPCVQV-----NISGADMSYVCKPV-----EYSISVSNP-----GDLVL	356
Db	595	ASASAG-PPLISGTCGLTVGDGHHKHNEBSHHDCRECYCLUNGREMCALITCPVPACNGPTI	653
Qy	357	H-----DVIQTDTLGSYTVLEAPGEICCNKVWRI-----KE	390
Db	654	HPQCPCSCADDFVQRPCLSTPSICHAPGGEYFVEGETWNIIDSTQCTCHSGRVLCEPE	713
Qy	391	MCPEGTQLKLVVAQVPGRTNOVAVTSESNGCTCSCAETTHWKGLAATHMCVLDTN	450
Db	714	VCP-----PLUCQNSR-----TDSCCPQCTQPPRPSLSRNNSVPNCKNDEG	758
Qy	451	DPICVGVNTVVRICVT	466
Db	759	DIFLAESKMPDVCTS	774

RESULT 14

RESOL 14
US-10-121-063-142
; Sequence 142, Application US/10121063

```

1  GENERAL INFORMATION:
2  APPLICANT: Baker, Kevin P.
3  APPLICANT: Beresini, Maureen
4  APPLICANT: DeGeorge, Laura
5  APPLICANT: Desnoyers, Luc
6  APPLICANT: Filvaroff, Ellen
7  APPLICANT: Gao, Wei-Qiang
8  APPLICANT: Gerlitsen, Mary E.
9  APPLICANT: Goddard, Audrey
10 APPLICANT: Godowski, Paul J.
11 APPLICANT: Gurney, Austin L.
12 APPLICANT: Sherwood, Steven
13 APPLICANT: Smith, Victoria
14 APPLICANT: Stewart, Timothy A.
15 APPLICANT: Tumas, Daniel
16 APPLICANT: Watanabe, Colin K
17 APPLICANT: Wood, William
18 APPLICANT: Zhang, Zemin
19 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
20 TITLE OF INVENTION: ACIDS ENCODING THE SAME
21 FILE REFERENCE: P3330R1G19
22 CURRENT APPLICATION NUMBER: US/10/121,063
23 CURRENT FILING DATE: 2002-04-12
24 Prior Application removed - See File Wrapper or Palm
25 NUMBER OF SEQ ID NOS: 550

```

```

Query Match      3.9%  Score 114.5;  DB 6;  Length 1036;
Best Local Similarity 19.2%;  Pres. No. 0.75;
Matches 95;  Conservative 45;  Mismatches 155;  Indels 201;  Gaps 24;
QY 75 CDKEFYPCDEGRCQVEAQSCYGRLYSVK-----VNDDCNVEICQ 116

```

RESULT 15

```

RESULTS 13
US-10-123-108-142
> Sequence 142, Application US/10123108
> GENERAL INFORMATION:
> APPLICANT: Baker, Kevin P.
> APPLICANT: Beresini, Maureen
> APPLICANT: Deforge, Laura
> APPLICANT: Desnoyers, Luc
> APPLICANT: Filvaroff, Ellen
> APPLICANT: Gao, Wei-Qiang
> APPLICANT: Gerritsen, Mary E.
> APPLICANT: Goddard, Audrey
> APPLICANT: Godowski, Paul J.
> APPLICANT: Gurney, Austin L.
> APPLICANT: Sherwood, Steven
> APPLICANT: Smith, Victoria
> APPLICANT: Stewart, Timothy A.
> APPLICANT: Tumas, Daniel
> APPLICANT: Watanabe, Colin K
> APPLICANT: Wood, William
> APPLICANT: Zhang, Zemin
> TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
> ACIDS ENCODING THE SAME
> FILE REFERENCE: P330R1C36
> CURRENT APPLICATION NUMBER: US/10/123,108
> CURRENT FILING DATE: 2002-04-15
> PRIOR APPLICATION NUMBER: 60/049911
> PRIOR FILING DATE: 1997-06-18
> PRIOR APPLICATION NUMBER: 60/056974
> PRIOR FILING DATE: 1997-08-26
> PRIOR APPLICATION NUMBER: 60/059113
> PRIOR FILING DATE: 1997-09-17
> PRIOR APPLICATION NUMBER: 60/059115
> PRIOR FILING DATE: 1997-09-17
> PRIOR APPLICATION NUMBER: 60/059117
> PRIOR FILING DATE: 1997-09-17
> PRIOR APPLICATION NUMBER: 60/059122

```



```
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090538
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07

Query Match      3.9%; Score 114.5; DB 6; Length 1036;
Best Local Similarity 19.2%; Pred. No. 0.75; 155; Indels 201; Gaps 24;
Matches 99; Conservative 45; Mismatches 45;

QY 75 CDKEFYPCERGROPEAQESCYRLYSVK-----VNDCCNVEICQ 116
   |:::| | | | | | | | | | | | | | | | | | | | | |
Db 376 CERYVVP--EGECCPV-----CEDPVFPFNNPAGCYANGLILAHGDRNRDDCT--FCQ 425
   |:::| | | | | | | | | | | | | | | | | | | | | |

QY 117 SVPEYATVSPYPIEILAIKKDCVDVVIQTQ-----QLPCEAEFVSSDP---ETTPTSD 167
   |:::| | | | | | | | | | | | | | | | | | | | | |
Db 426 CVN-----GERHCYATVCGGTCTNPVKVPGCECPVCEPTIITVDPPAC 469
   |:::| | | | | | | | | | | | | | | | | | | | | |

QY 168 GKLK-----WKIDRLCAGDKCKITVW-----KPLKEGC--C---FTAAT 202
   |:::| | | | | | | | | | | | | | | | | | | | | |
Db 470 GELSNCLTSGKDCINGFKRDHNG---CRTQCINTEELCSERKQGCTLNCPEGLTDAQ 525
   |:::| | | | | | | | | | | | | | | | | | | | | |

QY 203 VCACPELRSYTKGQPAICIK-----OEGPD-CACLRCP--VCYKIEVNTGSATA 250
   |:::| | | | | | | | | | | | | | | | | | | | | |
Db 526 NCEICECRPRPKCRPLICDKYCPGLLKNKHGCDICRCKKPELSCKSI-----575
   |:::| | | | | | | | | | | | | | | | | | | | | |

QY 251 RNVTVDPNPDGYSHASGQRLVSFNLGDMRPGDKKVFVEFCPORRGQITNVATVYCGG 310
   |:::| | | | | | | | | | | | | | | | | | | | | |
Db 576 -----CPLGFOODSHGCLICKCRE 594
   |:::| | | | | | | | | | | | | | | | | | | | | |

QY 311 HKCSANTVTVNPEPCVQV-----NIGADWSYVCKPV-----EYSISVSNP-----GDLVL 356
   |:::| | | | | | | | | | | | | | | | | | | | | |
Db 595 ASASAG-PPILSGTCLTVDGHHKNEESWHHDGCRECYCLNGREMCALITCPVPACGNPTI 653
   |:::| | | | | | | | | | | | | | | | | | | | | |

QY 357 H-----DVVIQDTLPSTVLEAPGGEICGNKVWRI-----KE 390
   |:::| | | | | | | | | | | | | | | | | | | | | |
Db 654 HPGCCPCSCADDFVQKPELSTPSICHAPGGEYFVEGETWNIDSCOTCTCHSGRVLCEETE 713
   |:::| | | | | | | | | | | | | | | | | | | | | |

QY 391 MCPGETLQFKLVKQVPGFRFTNOAVTSESNCGTCTSCAETTHKGLAATHMVCVLDTN 450
   |:::| | | | | | | | | | | | | | | | | | | | | |
Db 714 VCP-----PLLCQNPFR-----TQDSCCPOCTQDPRPSLRNNSVNPYCNDEG 758
   |:::| | | | | | | | | | | | | | | | | | | | | |

QY 451 DPICVGENTVYRICVT 466
   |:::| | | | | | | | | | | | | | | | | | | | | |
Db 759 DIFLAESWKPVDVCTS 774
   |:::| | | | | | | | | | | | | | | | | | | | | |
```

Search completed: May 25, 2002, 22:22:16
Job time: 463 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2002, 22:21:04 : Search time 35.96 Seconds
(without alignments)
1717.381 Million cell updates/sec

Title: US-09-523-647-2

Perfect score: 556

Sequence: 1 MSKLIRRVTVLTALTSNASCFASGGIEAAVBSLITKIVASAEKPAVPMTAKVRLVR 556

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size: 12

Total number of hits satisfying chosen parameters: 24

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: A_Geneseq_032802.*
1: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

length of match
SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	556	100.0	556	AA1980	Chlamydia pneumoniae
2	556	100.0	556	AA1981	A 50 kDa cysteine-
3	25	4.5	553	AA1978	Chlamydia trachoma
4	16	2.9	16	AA1979	Chlamydia pneumoniae
5	16	2.9	16	AA1980	C. pneumoniae OMCB
6	16	2.9	16	AA1981	Chlamydia pneumoniae
7	16	2.9	16	AA1982	Chlamydia trachoma
8	16	2.9	20	AA1983	C. trachomatis OMC
9	16	2.9	20	AA1984	Chlamydia trachoma
10	13	2.3	109	AA1985	Chlamydia Crp gene
11	12	2.2	12	AA1986	C. trachomatis OMC

12	12	2.2	12	22	AAG83244	Chlamydia trachoma
13	12	2.2	16	21	AA13675	C. trachomatis OMC
14	12	2.2	16	22	AAG83243	Chlamydia trachoma
15	12	2.2	20	21	AA13665	C. trachomatis OMC
16	12	2.2	20	21	AA13666	C. trachomatis OMC
17	12	2.2	20	21	AA13670	C. trachomatis OMC
18	12	2.2	20	21	AA13671	C. trachomatis OMC
19	12	2.2	20	21	AA13673	Chlamydia trachoma
20	12	2.2	20	22	AAG83233	Chla
21	12	2.2	20	22	AAG83234	Chla
22	12	2.2	20	22	AAG83238	Chla
23	12	2.2	20	22	AAG83239	Chla
24	12	2.2	20	22	AAG83241	Chla

ALIGNMENTS

RESULT 1

AA1980
ID AAY35178 standard; Protein; 556 AA.

XX AC AAY35178;

XX DT 13-SEP-1999 (first entry)

XX DE Chlamydia pneumoniae transmembrane protein sequence.

XX KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
XX KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
XX KW vaccine; neutralising epitope.

XX OS Chlamydia pneumoniae.

XX PN W09927105-A2.

XX PD 03-JUN-1999.

XX PF 20-NOV-1998; 98WO-1B01890.

XX PR 04-NOV-1998; 98US-0107078.

XX PR 21-NOV-1997; 97FR-0014673.

XX PA (GEST) GENSET.

XX PI Griffais R;

XX WP; 1999-357842/30.

XX PT Genome sequence of Chlamydia pneumoniae

XX PS Page 1037-1038; Disclosure; 1912pp; English.

XX CC AAY34584-Y35879 represent the proteins encoded by all the open reading
frames in the complete genome (see AAY34584-Y35879) of Chlamydia pneumoniae.
XX CC C. pneumoniae causes respiratory disease such as pneumonia and
XX CC bronchitis and is thought to be a contributing factor in heart
XX CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
XX CC nodosum or pharyngitis. The polypeptides encoded by the open reading
XX CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
XX CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
XX CC nucleotide sequences can also be used as immunogenic compositions,
XX CC especially where the vector directs the expression of a neutralising
XX CC epitope of C. pneumoniae.

XX SQ Sequence 556 AA;

Query Match 100.0%; Score 556; DB 20; Length 556;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKLIRRVTVLTALTSNASCFASGGIEAAVBSLITKIVASAEKPAVPMTAKVRLVR 60

Db 1 mskllrrvvvtvltasmascfsggieaavaeslltklvasaetkppapvmtakkrvlvr 60
 Qy 61 RNKQPVESKSGAFCDKEFYPCBEGRCQPVFAQQESCYGRLYSVKVNDDCNVEICQSVPE 120
 Db 61 rnkqpvdksgatgatkdfeypceegrcqpvfaadqescygrlysvkvnddcnveicqsvpe 120
 Qy 121 YATVGSPPYPIETLAIKGGKDCVVDVITQOLPCEAEFVSSDPETPTSDGKLVWKIDRLGAG 180
 Db 121 yatvgsppyleiailgkdkcdvvdvtqglpceaeefvssdpettptsdgklvkwidrlgag 180
 Qy 181 DKCKITVWVKPLKEGCGCFTAAATVACAPELRSYTKCGQPAICIKQSGPDCACLRCPVCYKI 240
 Db 181 dkckitvwvplkegcgcftaatvaccapelsytkcgqpaicikqsgpdcacrlcpvcyki 240
 Qy 241 EVVNTGSAIARNVTVDNVPDGYSHASQORVLSFNLGDMRPGDKKVFVFCPQRRGQIT 300
 Db 241 evvntgsaiaarnvtvdnvpdgyshaagqrvisfnlgdmrpgdkkfvfvcpcqrrgqit 300
 Qy 301 NVATVYCGGHKCSANVTTVVNEPCVQVNIISGADWSYCKPVEYSISVSNPGDLVLHDVV 360
 Db 301 nvatvycgghkcsanvttvvnepcvqvnisgadwsyckpveyisvsnpgdlvlhdvv 360
 Qy 361 IODTLPSTGVTLEAPGGEICCNKVMRIKEMCPGETLQFKLVYKAOVPGRTNQAVTSE 420
 Db 361 iqdtpstgvtvleapggeiccnkvwrkicmcpgetlqfklvkaqvpggrtngvavtse 420
 Qy 421 SNGCTCTCAETTHWKGLAATHMCLVDTNDPICVGENTVYRICVTRNGSAEDTNVSLIL 480
 Db 421 sngctctcaetthwkglaathmclvldtndpicvgentvyricevtrngsaedtnvslil 480
 Qy 481 KFSKELQPIASSGPTKGTISGNTVVDALPKLGSKEVSFEVSTLKGIAPGDARGEAILSS 540
 Db 481 kfskelqpiassgptktisgntvvfdalpklgskesvefsvtlkgiapgdargeailss 540
 Qy 541 DTLTSPVSDTENTHVV 556
 Db 541 dltspvsdtenhvv 556

RESULT 2

AA18820
 ID AAB18820 standard; Protein; 556 AA.

AC AAB18820;

DT 22-JAN-2001 (first entry)

XX A 60 kda cysteine-rich membrane protein of Chlamydia pneumoniae.

XX Cysteine-rich membrane protein; Chlamydia infection; bronchitis;
 KW community acquired pneumonia; upper respiratory tract infection; vaccine;
 KW sinusitis.

XX Chlamydia pneumoniae.

XX WO200055326-A1.

XX 21-SEP-2000.

XX 09-MAR-2000; 2000WO-CA00240.

XX 12-MAR-1999; 99US-0123966.

XX (AVET) AVENTIS PASTEUR LTD.

XX Murdin AD, Oomen RP, Wang J, Dunn P;

XX WPI; 2000-618918/59.

XX N-PSDB; AAB75901.

XX New polynucleotides encoding a 60kda cysteine-rich membrane protein
 PT from Chlamydia, useful as a vaccine for preventing and treating

PT Chlamydia infection in mammals -
 XX Claim 16; Fig 1; 77pp; English.

CC The present sequence represents a Chlamydia 60 kda cysteine-rich membrane
 CC protein. The membrane-rich polynucleotide and polypeptide are useful
 CC for preventing or treating Chlamydia infection, such as community
 CC pneumonia, upper respiratory tract infections, bronchitis and sinusitis.
 CC They are also useful for diagnosing Chlamydia infection by assaying a
 CC body fluid of a mammal. The polypeptide is useful for vaccine
 CC production.

XX Sequence 556 AA;

Query Match 100.0%; Score 556; DB 21; Length 556;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKLIRRVVTVLTALTSMSCFASGGIEAAVAESLITTKIVASAETKPPAPVMTAKKRVLR 60

Db 1 mskllrrvvvtvltasmascfsggieaavaeslltklvasaetkppapvmtakkrvlvr 60

Qy 61 RNKQPVESKSGAFCDKEFYPCBEGRCQPVFAQQESCYGRLYSVKVNDDCNVEICQSVPE 120

Db 61 rnkqpvdksgatgatkdfeypceegrcqpvfaadqescygrlysvkvnddcnveicqsvpe 120

Qy 121 YATVGSPPYPIETLAIKGGKDCVVDVITQOLPCEAEFVSSDPETPTSDGKLVWKIDRLGAG 180

Db 121 yatvgsppyleiailgkdkcdvvdvtqglpceaeefvssdpettptsdgklvkwidrlgag 180

Qy 181 DKCKITVWVKPLKEGCGCFTAAATVACAPELRSYTKCGQPAICIKQSGPDCACLRCPVCYKI 240

Db 181 dkckitvwvplkegcgcftaatvaccapelsytkcgqpaicikqsgpdcacrlcpvcyki 240

Qy 241 EVVNTGSAIARNVTVDNVPDGYSHASQORVLSFNLGDMRPGDKKVFVFCPQRRGQIT 300

Db 241 evvntgsaiaarnvtvdnvpdgyshaagqrvisfnlgdmrpgdkkfvfvcpcqrrgqit 300

Qy 301 NVATVYCGGHKCSANVTTVVNEPCVQVNIISGADWSYCKPVEYSISVSNPGDLVLHDVV 360

Db 301 nvatvycgghkcsanvttvvnepcvqvnisgadwsyckpveyisvsnpgdlvlhdvv 360

Qy 361 IODTLPSTGVTLEAPGGEICCNKVMRIKEMCPGETLQFKLVYKAOVPGRTNQAVTSE 420

Db 361 iqdtpstgvtvleapggeiccnkvwrkicmcpgetlqfklvkaqvpggrtngvavtse 420

Qy 421 SNGCTCTCAETTHWKGLAATHMCLVDTNDPICVGENTVYRICVTRNGSAEDTNVSLIL 480

Db 421 sngctctcaetthwkglaathmclvldtndpicvgentvyricevtrngsaedtnvslil 480

Qy 481 KFSKELQPIASSGPTKGTISGNTVVDALPKLGSKEVSFEVSTLKGIAPGDARGEAILSS 540

Db 481 kfskelqpiassgptktisgntvvfdalpklgskesvefsvtlkgiapgdargeailss 540

Qy 541 DTLTSPVSDTENTHVV 556

Db 541 dltspvsdtenhvv 556

RESULT 3

AA18820
 ID AAB18820 standard; Protein; 553 AA.

AC AAB18820;

DT 07-OCT-1999 (first entry)

XX Chlamydia trachomatis cellular envelope protein.

XX Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
 KW paratrachoma; inclusion conjunctivitis; genital disease; perilepithelitis;
 KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;

KW bartholinitis; pneumopathy; venereal lymphogranulomatosis.

OS Chlamydia trachomatis.

PN WO9928475-A2.

XX 10-JUN-1999.

XX 27-NOV-1998; 98WO-TB01939.

XX 04-NOV-1998; 98US-0107077.

XX 28-NOV-1997; 97FR-0015041.

XX 17-DEC-1997; 97FR-0016034.

XX (GEST) GENSET.

XX Griffais R;

XX WPI; 1999-371125/31.

XX Genome sequence of Chlamydia trachomatis

XX Disclosure; Page 1371-1372; 1755pp; English.

XX AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
CC of Chlamydia trachomatis (see AA201425). The polypeptides can be used as
CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
CC can also be used to control growth of the microorganism. Chlamydia
CC trachomatis is responsible for a large number of diseases, e.g. eye
CC diseases such as conventional trachoma, nonendemic trachoma,
CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
CC perihepatitis, bartholinitis; pneumopathy in breast feeding infants;
CC and venereal lymphogranulomatosis. The polypeptides of the invention
CC may be of use in treating these diseases.

XX Sequence 553 AA;

Query Match 4.5%; Score 25; DB 20; Length 553;

Best Local Similarity 100.0%; Pred. No. 9.4e-17;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 KITVWKPLKEGCGCFTAAATVCAPE 208

Db 181 kitvwwkplkegcgcftaatvcaape 205

RESULT 4

AAI42728

ID AAI42728 standard; peptide; 16 AA.

XX AC AAY42728;

XX 20-DEC-1999 (first entry)

XX Chlamydia pneumoniae outer membrane protein-derived peptide.

XX Heart disease; inflammatory; autoimmune; cardiomyopathy; myosin;

XX Chlamydia; induction; vaccine.

XX Synthetic.

XX Chlamydia pneumoniae.

XX Key Location/Qualifiers

FT Modified-site 1

FT /note= "optionally N-terminally acetylated"

FT Misc-difference 16

FT /note= "Encoded by GTC"

XX US5962636-A.

XX 05-OCT-1999.

XX 12-AUG-1998; 98US-0133774.

XX 12-AUG-1998; 98US-0133774.

XX (AMGE-) AMGEN CANADA INC.

XX Bachmaier K, Hessel AJ, Penninger JM, Neu N;

XX WPI; 1999-589735/50.

XX N-PSDB; AA228184.

XX Peptides that induce or suppress inflammatory cardiomyopathy

XX Claim 1; Column 19; 17pp; English.

XX This sequence represents Chlamydia pneumoniae outer membrane protein-
CC derived peptide, which induces inflammatory cardiomyopathy (ICM) via
CC an autoimmune response in mice immunised with it. It contains an amino
CC acid sequence motif Maxxs (AA42722) which appears to be required for
CC the induction of this disease. The motif was originally identified in
CC the sequence of murine alpha myosin heavy chain-derived peptide,
CC M7A-alpha, when it was compared with a peptide derived from a homologous
CC region of the murine beta myosin heavy chain, M7A-beta (AA42724) which
CC did not cause the disease. Several peptide fragments containing the
CC motif were identified from a database and were found to be fragments of
CC cysteine rich outer membrane proteins from various species of Chlamydia.
CC These peptides also induced ICM, indicating that infection with
CC Chlamydia may be involved in the development of ICM. Inflammatory
CC cardiomyopathy peptides are used to determine the risk of ICM by
CC incubation with a subject's T cells and measuring the degree of
CC proliferation (an increased degree being indicative of risk) or to raise
CC specific antibodies which can be used therapeutically and for the
CC detection of Chlamydia. Such peptides can also be used with an adjuvant
CC and an excipient in a vaccine for decreasing ICM.

XX Sequence 16 AA;

Query Match 2.9%; Score 16; DB 20; Length 16;

Best Local Similarity 100.0%; Pred. No. 8.4e-09;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GIEAAVAESLITKIVA 40

Db 1 gieaavaeslitriva 16

RESULT 5

AAI13677

ID AAI13677 standard; Peptide; 16 AA.

XX AC AAI13677;

XX 02-FEB-2001 (first entry)

XX C. pneumoniae OMCB peptide 185-198.

XX Chlamydial infection; sexually transmitted disease;

XX pelvic inflammatory disease; PID; tubal obstruction; infertility;

XX trachoma; blindness; acute respiratory tract infection;

XX atherosclerosis; coronary heart disease; antibacterial.

XX Chlamydia pneumoniae.

XX WO200034483-A2.

XX 15-JUN-2000.

XX 08-DEC-1999; 99WO-US29012.

XX 08-DEC-1998; 98US-0208277.

XX 08-APR-1999; 99US-0288594.

```

PR 01-OCT-1999; 99US-0410568.
XX 22-OCT-1999; 99US-0426571.
XX (CORI-) CORIXA CORP.
XX
PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX
XX WPI: 2000-431303/37.
XX
XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection
PT comprises immunogenic portion of Chlamydia antigen, which comprises
PT amino acid sequence encoded by polynucleotide sequence -
XX
XX Disclosure; Page 233; 256pp; English.
XX
XX The present invention relates to new nucleic acid sequences and the
CC proteins encoded by the nucleic acid sequences. The encoded proteins
CC comprise an immunogenic portion of a Chlamydia antigen. The encoded
CC proteins are useful for the serodiagnosis and treatment of Chlamydia
CC infection. Chlamydiae are intracellular bacterial pathogens that are
CC responsible for a wide variety of human infections. C. trachomatis
CC infection is one of the most common sexually transmitted diseases and can
CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
CC and infertility. Trachoma due to ocular infection with C. trachomatis is
CC the leading cause of preventable blindness worldwide. C. pneumoniae is a
CC major cause of acute respiratory tract infections in humans and is also
CC thought to play a role in the pathogenesis of atherosclerosis and
CC coronary heart disease. The present sequence is a protein isolated in the
XX present invention.
XX
SQ Sequence 16 AA;

Query Match 2.9%; Score 16; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.4e-09;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GDKCKITVWKPLKEG 195
DB 1 gdkckitvwwkplkeg 16
|||||
1 gdkckitvwwkplkeg 16

RESULT 6
AAY83817
ID AAY83817 standard; Protein; 16 AA.
AC AAY83817;
XX
XX 21-JUN-2000 (first entry)
XX
XX Chlamydia pneumoniae Cys-rich OMP peptide.
DE
XX
XX Cardiant; murine alpha myosin heavy chain; inflammatory myocarditis;
KW autoimmune inflammatory cardiomyopathy; Chlamydia; antibody; vaccine;
KW hybridization probe.
XX
XX Chlamydia pneumoniae.
OS
XX
XX US6034230-A.
PN
XX
XX 07-MAR-2000.
PD
XX
XX 03-MAY-1999; 99US-0303862.
XX
XX 12-AUG-1998; 98US-0133774.
XX
XX (AMGE-) AMGEN CANADA INC.
PA
XX
XX Neu N, Penninger JM, Bachmaier K, Hessel AJ;
PI
XX WPI: 2000-255712/22.
DR
XX N-PSDB: AAZ99167.
XX

PT DNA molecules encoding novel myocardial peptides used for inhibiting
PT and inducing inflammatory cardiomyopathy in vivo -
XX
XX Example 1; Column 18; 17pp; English.
XX
XX This sequence represents a fragment of the Chlamydia pneumoniae Cys-rich
CC outer membrane protein (OMP). The sequence was obtained by homology
CC searches of the PIR database using the murine M7A-alpha/beta peptide
CC consensus sequence (Y83813) as the query. The screen isolated the
CC peptides Y83814-Y83819 and their corresponding coding sequences
CC 259164-259169. The invention relates to the isolation of sequences
CC coding for peptide sequences derived from bacteria and viruses which may
CC cause inflammatory cardiomyopathy. The peptide sequences are searched
CC based on the sequence of the M7a peptides derived from the murine alpha
CC myosin heavy chain polypeptide. The peptides encoded by the DNAs are
CC used, alone or in conjunction with other therapeutics, for inducing or
CC inhibiting inflammatory cardiomyopathy in vivo, where the cardiomyopathy
CC is autoimmune inflammatory cardiomyopathy, and inflammatory
CC cardiomyopathy caused by Chlamydia or other bacterial or viral infections
CC that cause inflammatory myocarditis in a mammal. Antibodies against the
CC peptides and the peptides themselves are used for measuring the risk of
CC inflammatory cardiomyopathy in a mammal. The peptides may also be used
CC in vaccines. Nucleic acids encoding the peptides may be used as
CC hybridization probes, e.g. in diagnostic assays to test for the
XX presence of Chlamydia DNA.
XX
SQ Sequence 16 AA;

Query Match 2.9%; Score 16; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.4e-09;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GIEAAVAESLITKIYA 40
DB 1 gieaavaeslilkiya 16
|||||
1 gieaavaeslilkiya 16

RESULT 7
AAG83245
ID AAG83245 standard; Peptide; 16 AA.
AC AAG83245;
XX
XX 05-SEP-2001 (first entry)
XX
XX Chlamydia trachomatis OMCB peptide 185-198.
DE
XX
XX Chlamydia; vaccine; infection; fusion protein; antigen;
KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
KW acute respiratory tract infection; Capi; C7529; OMCB;
KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.
XX
XX Chlamydia trachomatis.
OS
XX
XX WO200140474-A2.
PN
XX
XX 07-JUN-2001.
PD
XX
XX 04-DEC-2000; 2000WO-US32919.
XX
XX 03-DEC-1999; 99US-0454684.
XX
XX 19-APR-2000; 2000US-0556877.
XX
XX 20-JUN-2000; 2000US-0598419.
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;
PI
XX WPI: 2001-374831/39.
XX
XX Chlamydia polypeptides and fusion proteins useful for preventing pelvic
PT

```

PT inflammatory disease, trachoma, acute respiratory tract infections,
PT atherosclerosis and heart disease -
XX
XX Claim 66; Page 239; 295pp; English.
XX
CC The present peptide is provided in a specification relating to
CC compounds and methods for the treatment and diagnosis of Chlamydia
CC infection. The compounds provided include polypeptides and fusion
CC proteins comprising immunogenic portions of Chlamydia antigens
CC and DNA sequences encoding such polypeptides. They are useful for
CC vaccinating against chlamydial infection, which causes pelvic
CC inflammatory disease, trachoma, acute respiratory tract infections,
CC atherosclerosis and heart disease.
XX
XX Sequence 16 AA;
SQ
Query Match 2.9%; Score 16; DB 22; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.4e-09; Mismatches 0; Indels 0; Gaps 0;
Matches 16; Conservative 0;
CY 180 GDKCKITVWVKPLKEG 195
DB 1 gdkckitvwvkplkeg 16
RESULT 8
AAB13674
ID AAB13674 standard; Peptide; 20 AA.
XX
AC AAB13674;
XX
XX 02-FEB-2001 (first entry)
XX
DE C. trachomatis OMCB peptide 171-190.
XX
XX Chlamydial infection; sexually transmitted disease;
XX pelvic inflammatory disease; PID; tubal obstruction; infertility;
XX trachoma; blindness; acute respiratory tract infection;
XX atherosclerosis; coronary heart disease; antibacterial.
XX
XX Chlamydia trachomatis.
XX
XX WO2000034483-A2.
XX
XX 15-JUN-2000.
XX
XX 08-DEC-1999; 99WO-US29012.
XX
XX 08-DEC-1998; 98US-0208277.
XX 08-APR-1999; 99US-0288594.
XX 01-OCT-1999; 99US-0410568.
XX 22-OCT-1999; 99US-0426571.
XX
XX (CORI-) CORIXA CORP.
XX
XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX WPI; 2000-431303/37.
XX
XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection
XX comprises immunogenic portion of Chlamydia antigen, which comprises
XX amino acid sequence encoded by polynucleotide sequence -
XX
XX Disclosure; Page 233; 256pp; English.
XX
XX The present invention relates to new nucleic acid sequences and the
XX proteins encoded by the nucleic acid sequences. The encoded proteins
XX comprise an immunogenic portion of a Chlamydia antigen. The encoded
XX proteins are useful for the serodiagnosis and treatment of Chlamydia
XX infection. Chlamydiae are intracellular bacterial pathogens that are
XX responsible for a wide variety of human infections. C. trachomatis
XX infection is one of the most common sexually transmitted diseases and can

CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
CC and infertility. Trachoma due to ocular infection with C. trachomatis is
CC the leading cause of preventable blindness worldwide. C. pneumonia is a
CC major cause of acute respiratory tract infections in humans and is also
CC thought to play a role in the pathogenesis of atherosclerosis and
CC coronary heart disease. The present sequence is a protein isolated in the
CC present invention.
XX
XX Sequence 20 AA;
SQ
Query Match 2.9%; Score 16; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 1e-08; Mismatches 0; Indels 0; Gaps 0;
Matches 16; Conservative 0;
CY 184 KITVWVKPLKEGCGFT 199
DB 5 kitvwvkplkegcgft 20
RESULT 9
AAG83242
ID AAG83242 standard; Peptide; 20 AA.
XX
XX AAG83242;
XX
XX 05-SEP-2001 (first entry)
XX
XX Chlamydia trachomatis OMCB peptide 171-190.
XX
XX Chlamydia; vaccine; infection; fusion protein; antigen;
XX pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
XX acute respiratory tract infection; Capl; CF529; OMCB; thiol specific antioxidant; TSA.
XX polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.
XX
XX Chlamydia trachomatis.
XX
XX WO200140474-A2.
XX
XX 07-JUN-2001.
XX
XX 04-DEC-2000; 2000WO-US32919.
XX
XX 03-DEC-1999; 99US-0454684.
XX 19-APR-2000; 2000US-0556877.
XX 20-JUN-2000; 2000US-0598419.
XX
XX (CORI-) CORIXA CORP.
XX
XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;
XX WPI; 2001-374831/39.
XX
XX Chlamydia polypeptides and fusion proteins useful for preventing pelvic
XX inflammatory disease, trachoma, acute respiratory tract infections,
XX atherosclerosis and heart disease -
XX Claim 66; Page 239; 295pp; English.
XX
XX The present peptide is provided in a specification relating to
XX compounds and methods for the treatment and diagnosis of chlamydial
XX infection. The compounds provided include polypeptides and fusion
XX proteins comprising immunogenic portions of Chlamydia antigens
XX and DNA sequences encoding such polypeptides. They are useful for
XX vaccinating against chlamydial infection, which causes pelvic
XX inflammatory disease, trachoma, acute respiratory tract infections,
XX atherosclerosis and heart disease.
XX
XX Sequence 20 AA;
SQ
Query Match 2.9%; Score 16; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 1e-08; Mismatches 0; Indels 0; Gaps 0;
Matches 16; Conservative 0;

```

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 KITVWRPLKGGCET 199
Db 5 kitvwkplkegcft 20

RESULT 10
AAR07399
ID AAR07399 standard; protein; 109 AA.
AC AAR07399;
XX
XX 12-MAR-1991 (first entry)
XX Chlamydia CrP gene prod.
XX
XX PCR; DIANA; solid phase medical diagnosis.
XX
XX Chlamydia trachomatis.
XX
XX WO9011369-A.
XX
XX 04-OCT-1990.
XX
XX 15-MAR-1990; 90WO-EP00454.
XX
XX 22-MAR-1989; 89GB-0006642.
XX 22-MAR-1989; 89GB-0006641.
XX
XX (CEMU-) CEMU BIOTEKNIK.
XX
XX Uhlen M;
XX
XX WPI; 1990-320269/42.
XX N-FSDb; AAQ06236.
XX
XX Amplification of DNA by PCR - using distal DNA on a primer as a
XX handle for attachment to a solid support or a label, used in
XX diagnosis of medical conditions
XX
XX Disclosure; fig 19; 66pp; English.
XX
XX This is the product encoded by the CrP gene. The corresp. gene can
XX be detected utilising DIANA, an improved PCR method with 2 primer
XX pairs.
XX
XX Sequence 109 AA;

Query Match 2.3%; Score 13; DB 11; Length 109;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 GADWSYVCKPVEY 344
Db 30 gadwsyvcpkvey 42

RESULT 11
AAB13676
ID AAB13676 standard; Peptide; 12 AA.
XX
XX AAB13676;
XX
XX 02-FEB-2001 (first entry)
XX
XX C. trachomatis OMCB peptide 175-186.
XX
XX Chlamydial infection; sexually transmitted disease;
XX pelvic inflammatory disease; PID; tubal obstruction; infertility;
XX trachoma; blindness; acute respiratory tract infection;
XX atherosclerosis; coronary heart disease; antibacterial.

```

```

XX Chlamydia trachomatis.
OS
XX WO200034483-A2.
XX
XX 15-JUN-2000.
XX
XX 08-DEC-1999; 93WO-US29012.
XX
XX 08-DEC-1998; 98US-0208277.
XX 08-APR-1999; 99US-0288594.
XX 01-OCT-1999; 99US-0410568.
XX 22-OCT-1999; 99US-0426571.
XX
XX (CORI-) CORIXA CORP.
XX
XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX WPI; 2000-431303/37.
XX
XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection
XX comprises immunogenic portion of Chlamydia antigen, which comprises
XX amino acid sequence encoded by polynucleotide sequence .
XX
XX Disclosure; Page 233; 256pp; English.
XX
XX The present invention relates to new nucleic acid sequences and the
XX proteins encoded by the nucleic acid sequences. The encoded proteins
XX comprise an immunogenic portion of a Chlamydia antigen. The encoded
XX proteins are useful for the serodiagnosis and treatment of Chlamydia
XX infection. Chlamydiae are intracellular bacterial pathogens that are
XX responsible for a wide variety of human infections. C. trachomatis
XX infection is one of the most common sexually transmitted diseases and can
XX lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
XX and infertility. Trachoma due to ocular infection with C. trachomatis is
XX the leading cause of preventable blindness worldwide. C. pneumonia is a
XX major cause of acute respiratory tract infections in humans and is also
XX thought to play a role in the pathogenesis of atherosclerosis and
XX coronary heart disease. The present sequence is a protein isolated in the
XX present invention.
XX
XX Sequence 12 AA;

Query Match 2.2%; Score 12; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 8.5e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 KITVWRPLKKEG 195
Db 1 kitvwkplkeg 12

RESULT 12
AAG83244
ID AAG83244 standard; Peptide; 12 AA.
XX
XX AAG83244;
XX
XX 05-SEP-2001 (first entry)
XX
XX Chlamydia trachomatis OMCB peptide 175-186.
XX
XX Chlamydia; vaccine; infection; fusion protein; antigen;
XX pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
XX acute respiratory tract infection; CapJ; CT529; OMCB;
XX polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.
XX
XX Chlamydia trachomatis.
OS
XX WO200140474-A2.
XX
XX 07-JUN-2001.

```

```
XX PF 04-DEC-2000; 2000WO-US32919.
XX PR 03-DEC-1999; 99US-0454684.
XX PR 19-APR-2000; 2000US-0556877.
XX PR 20-JUN-2000; 2000US-0598419.
XX PA (CORI-) CORIXA CORP.
XX PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;
XX DB WPI; 2001-374831/39.
XX PA Chlamydia polypeptides and fusion proteins useful for preventing pelvic
XX PR inflammatory disease, trachoma, acute respiratory tract infections,
XX PR atherosclerosis and heart disease -
XX PA Claim 66; Page 239; 295pp; English.
XX CC The present peptide is provided in a specification relating to
XX CC compounds and methods for the treatment and diagnosis of chlamydial
XX CC infection. The compounds provided include polypeptides of chlamydial
XX CC proteins comprising immunogenic portions of Chlamydia antigens
XX CC and DNA sequences encoding such polypeptides. They are useful for
XX CC vaccinating against chlamydial infection, which causes pelvic
XX CC inflammatory disease, trachoma, acute respiratory tract infections,
XX CC atherosclerosis and heart disease.
XX SQ Sequence 12 AA;
XX DB Query Match 2.2%; Score 12; DB 22; Length 12;
XX DB Best Local Similarity 100.0%; Pred. No. 8.5e-05;
XX DB Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 184 KITVWVKPLKEG 195
XX DB | | | | | | | | | |
XX AC 1 kitvwvkplkeg 12
XX DT 02-FEB-2001 (first entry)
XX DE C. trachomatis OMCB peptide 171-186.
XX KW Chlamydial infection; sexually transmitted disease;
XX KW pelvic inflammatory disease; PID; tubal obstruction; infertility;
XX KW trachoma; blindness; acute respiratory tract infection;
XX KW atherosclerosis; coronary heart disease; antibacterial.
XX OS Chlamydia trachomatis.
XX PN WO200034483-A2.
XX XX 15-JUN-2000.
XX PF 08-DEC-1999; 99WO-US29012.
XX PR 08-DEC-1998; 98US-0208277.
XX PR 08-APR-1999; 99US-0288594.
XX PR 01-OCT-1999; 99US-0410568.
XX PR 22-OCT-1999; 99US-0426571.
XX PA (CORI-) CORIXA CORP.
XX PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX XX WPI; 2000-431303/37.
XX DR

XX PF Isolated polypeptide for diagnosis and treatment of Chlamydia infection
XX PR comprises immunogenic portion of Chlamydia antigen, which comprises
XX PR amino acid sequence encoded by polynucleotide sequence -
XX PS Disclosure; Page 233; 256pp; English.
XX CC The present invention relates to new nucleic acid sequences and the
XX CC proteins encoded by the nucleic acid sequences. The encoded proteins
XX CC comprise an immunogenic portion of a Chlamydia antigen. The encoded
XX CC proteins are useful for the serodiagnosis and treatment of Chlamydia
XX CC infection. Chlamydiae are intracellular bacterial pathogens that are
XX CC responsible for a wide variety of human infections. C. trachomatis
XX CC infection is one of the most common sexually transmitted diseases and can
XX CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
XX CC and infertility. Trachoma due to ocular infection with C. trachomatis is
XX CC the leading cause of preventable blindness worldwide. C. pneumoniae is a
XX CC major cause of acute respiratory tract infections in humans and is also
XX CC thought to play a role in the pathogenesis of atherosclerosis and
XX CC coronary heart disease. The present sequence is a protein isolated in the
XX CC present invention.
XX SQ Sequence 16 AA;
XX DB Query Match 2.2%; Score 12; DB 21; Length 16;
XX DB Best Local Similarity 100.0%; Pred. No. 0.00011;
XX DB Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 184 KITVWVKPLKEG 195
XX DB | | | | | | | | | |
XX AC 5 kitvwvkplkeg 16
XX DT 05-SEP-2001 (first entry)
XX DE Chlamydia trachomatis OMCB peptide 171-186.
XX KW Chlamydia; vaccine; infection; fusion protein; antigen;
XX KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
XX KW acute respiratory tract infection; Capi; CF529; OMCB;
XX KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.
XX OS Chlamydia trachomatis.
XX PN WO200140474-A2.
XX XX 07-JUN-2001.
XX PF 04-DEC-2000; 2000WO-US32919.
XX PR 03-DEC-1999; 99US-0454684.
XX PR 19-APR-2000; 2000US-0556877.
XX PR 20-JUN-2000; 2000US-0598419.
XX PA (CORI-) CORIXA CORP.
XX PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;
XX XX WPI; 2001-374831/39.
XX PA Chlamydia polypeptides and fusion proteins useful for preventing pelvic
XX PR inflammatory disease, trachoma, acute respiratory tract infections,
XX PR atherosclerosis and heart disease -
XX XX Claim 66; Page 239; 295pp; English.
XX DR
```

CC The present peptide is provided in a specification relating to
CC compounds and methods for the treatment and diagnosis of chlamydial
CC infection. The compounds provided include polypeptides and fusion
CC proteins comprising immunogenic portions of Chlamydia antigens
CC and DNA sequences encoding such polypeptides. They are useful for
CC vaccinating against chlamydial infection, which causes pelvic
CC inflammatory disease, trachoma, acute respiratory tract infections,
CC atherosclerosis and heart disease.
XX Sequence 16 AA:
SQ

Query Match 2.28; Score 12; DB 22; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 KITVWVKPLKEG 195
|||||
DB 5 KITVWVKPLKEG 16
|||||

RESULT 15
AAB13665
ID AAB13665 standard; Peptide: 20 AA.
XX
AC AAB13665;
XX
DT 02-FEB-2001 (first entry)
XX
DE C. trachomatis OMCB peptide 128-147.
XX
KW Chlamydial infection; sexually transmitted disease;
KW pelvic inflammatory disease; PID; tubal obstruction; infertility;
KW trachoma; blindness; acute respiratory tract infection;
KW atherosclerosis; coronary heart disease; antibacterial.
XX
OS Chlamydia trachomatis.
XX
PN WQ200034483-A2.
XX
PD 15-JUN-2000.
XX
PF 08-DEC-1999; 99WO-US29012.
XX
PR 08-DEC-1998; 98US-0208277.
PR 08-APR-1999; 99US-0288594.
PR 01-OCT-1999; 99US-0410568.
PR 22-OCT-1999; 99US-0426571.
XX
PA (CORI-) CORIXA CORP.
XX
PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX
DR WPI; 2000-431303/37.

Isolated polypeptide for diagnosis and treatment of Chlamydia infection
comprises immunogenic portion of Chlamydia antigen, which comprises
amino acid sequence encoded by polynucleotide sequence
XX
PS Disclosure; Page 230; 256pp; English.
XX
CC The present invention relates to new nucleic acid sequences and the
CC proteins encoded by the nucleic acid sequences. The encoded proteins
CC comprise an immunogenic portion of a Chlamydia antigen. The encoded
CC proteins are useful for the serodiagnosis and treatment of Chlamydia
CC infection. Chlamidiae are intracellular bacterial pathogens that are
CC responsible for a wide variety of human infections. C. trachomatis
CC infection is one of the most common sexually transmitted diseases and can
CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
CC and infertility. Trachoma due to ocular infection with C. trachomatis is
CC the leading cause of preventable blindness worldwide. C. pneumoniae is a
CC major cause of acute respiratory tract infections in humans and is also
CC thought to play a role in the pathogenesis of atherosclerosis and

CC coronary heart disease. The present sequence is a protein isolated in the
CC present invention.
XX
SQ Sequence 20 AA:
Query Match 2.28; Score 12; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 ITQQLPCEAEFV 156
|||||
DB 9 ITQQLPCEAEFV 20
|||||

Search completed: May 25, 2002, 22:25:32
Job time: 268 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2002, 22:22:22 ; Search time 16.39 Seconds
(without alignments)
828.592 Million cell updates/sec

Title: US-09-523-647-2
Perfect score: 556
Sequence: 1 MSKLRRVTVLALTSMA...ILSSDTLTSPVSDTENTHVV 556

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 231628 seqs, 2442594 residues

Word size: 12
Total number of hits satisfying chosen parameters: 8

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	4.5	25	4 US-09-025-596-38	Sequence 38, Appl
2	20	3.6	20	4 US-09-025-596-41	Sequence 41, Appl
3	19	3.4	19	4 US-09-025-596-39	Sequence 39, Appl
4	16	2.9	16	2 US-09-133-774-7	Sequence 7, Appl
5	16	2.9	16	3 US-09-303-862-7	Sequence 7, Appl
6	16	2.9	16	4 US-09-025-596-43	Sequence 43, Appl
7	13	2.3	19	4 US-09-025-596-40	Sequence 40, Appl
8	13	2.3	109	1 US-08-477-270-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-09-025-596-38
; Sequence 38, Application US/09025596
; Patent No. 6340463
; GENERAL INFORMATION:
; APPLICANT: Mitchell, William M.
; APPLICANT: Stratton, Charles W.
; TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE
; TITLE OF INVENTION: SEQUENCES
; FILE REFERENCE: VDB98-01
; CURRENT APPLICATION NUMBER: US/09/025,596
; CURRENT FILING DATE: 1998-02-18
; EARLIER APPLICATION NUMBER: 08/911,593
; EARLIER FILING DATE: 1997-08-14

; EARLIER APPLICATION NUMBER: 60/023,921
; EARLIER FILING DATE: 1996-08-14
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 38
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-025-596-38

Query Match 4.5%; Score 25; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 2e-18;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 RNNQPVQKSRGAFCDKEFYPCEE 84
|||||
DB 1 RNNQPVQKSRGAFCDKEFYPCEE 25

RESULT 2
US-09-025-596-41
; Sequence 41, Application US/09025596
; Patent No. 6340463
; GENERAL INFORMATION:
; APPLICANT: Mitchell, William M.
; APPLICANT: Stratton, Charles W.
; TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE
; TITLE OF INVENTION: SEQUENCES
; FILE REFERENCE: VDB98-01
; CURRENT APPLICATION NUMBER: US/09/025,596
; CURRENT FILING DATE: 1998-02-18
; EARLIER APPLICATION NUMBER: 08/911,593
; EARLIER FILING DATE: 1997-08-14
; EARLIER APPLICATION NUMBER: 60/023,921
; EARLIER FILING DATE: 1996-08-14
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 41
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-025-596-41

Query Match 3.6%; Score 20; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.2e-13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 418 TSNSNCCTCTSCAFTTHWK 437
|||||
DB 1 TSNSNCCTCTSCAFTTHWK 20

RESULT 3
US-09-025-596-39
; Sequence 39, Application US/09025596
; Patent No. 6340463
; GENERAL INFORMATION:
; APPLICANT: Mitchell, William M.
; APPLICANT: Stratton, Charles W.
; TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE
; TITLE OF INVENTION: SEQUENCES
; FILE REFERENCE: VDB98-01
; CURRENT APPLICATION NUMBER: US/09/025,596
; CURRENT FILING DATE: 1998-02-18
; EARLIER APPLICATION NUMBER: 08/911,593
; EARLIER FILING DATE: 1997-08-14
; EARLIER APPLICATION NUMBER: 60/023,921
; EARLIER FILING DATE: 1996-08-14
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39

```
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-025-596-39

Query Match          3.4%; Score 19; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.2e-12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 278 DMRPGDKKVVTFECPQRR 296
Db 1 DMRPGDKKVVTFECPQRR 19

RESULT 4
US-09-133-774-7
; Sequence 7, Application US/09133774B
; Patent No. 5962636
; GENERAL INFORMATION:
; APPLICANT: Bachmaier, Kurt
; APPLICANT: Hessel, Andrew J.
; APPLICANT: Neu M.D., Nikolaus
; APPLICANT: Penninger, Josef M.
; TITLE OF INVENTION: No. 5962636el Peptides Capable of Modulating Inflammatory Heart
; FILE REFERENCE: A-536
; CURRENT APPLICATION NUMBER: US/09/133,774B
; CURRENT FILING DATE: 1998-08-12
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-133-774-7

Query Match          2.9%; Score 16; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 GIEAAVAESLITKIVA 40
Db 1 GIEAAVAESLITKIVA 16

RESULT 5
US-09-303-862-7
; Sequence 7, Application US/09303862
; Patent No. 6034230
; GENERAL INFORMATION:
; APPLICANT: Bachmaier, Kurt
; APPLICANT: Hessel, Andrew J.
; APPLICANT: Neu M.D., Nikolaus
; APPLICANT: Penninger, Josef M.
; TITLE OF INVENTION: No. 6034230el Peptides Capable of Modulating Inflammatory Heart
; FILE REFERENCE: A-536
; CURRENT APPLICATION NUMBER: US/09/303,862
; CURRENT FILING DATE: 1999-05-03
; EARLIER APPLICATION NUMBER: 09/133,774
; EARLIER FILING DATE: 1998-08-12
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-303-862-7

Query Match          2.9%; Score 16; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 GIEAAVAESLITKIVA 40
Db 1 GIEAAVAESLITKIVA 16

RESULT 6
US-09-025-596-43
; Sequence 43, Application US/09025596
; Patent No. 6340463
; GENERAL INFORMATION:
; APPLICANT: Mitchell, William M.
; APPLICANT: Stratton, Charles W.
; TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE
; FILE REFERENCE: VDB98-01
; CURRENT APPLICATION NUMBER: US/09/025,596
; CURRENT FILING DATE: 1998-02-18
; EARLIER APPLICATION NUMBER: 08/911,593
; EARLIER FILING DATE: 1997-08-14
; EARLIER APPLICATION NUMBER: 60/023,921
; EARLIER FILING DATE: 1996-08-14
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 43
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-025-596-43

Query Match          2.9%; Score 16; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 459 TVYRICVTNRGSAEDT 474
Db 1 TVYRICVTNRGSAEDT 16

RESULT 7
US-09-025-596-40
; Sequence 40, Application US/09025596
; Patent No. 6340463
; GENERAL INFORMATION:
; APPLICANT: Mitchell, William M.
; APPLICANT: Stratton, Charles W.
; TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE
; FILE REFERENCE: VDB98-01
; CURRENT APPLICATION NUMBER: US/09/025,596
; CURRENT FILING DATE: 1998-02-18
; EARLIER APPLICATION NUMBER: 08/911,593
; EARLIER FILING DATE: 1997-08-14
; EARLIER APPLICATION NUMBER: 60/023,921
; EARLIER FILING DATE: 1996-08-14
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 40
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-025-596-40

Query Match          2.3%; Score 13; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.2e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 157 SSDPETTPTSDGK 169
Db 1 SSDPETTPTSDGK 169
```


us-09-523-647-2.Olig.ra

Tue May 28 08:55:12 2002

DB 1 SSDPETPTSDGK 13

RESULT 8

US-08-477-270-20

Sequence 20, Application US/08477270

Patent No. 5629158

GENERAL INFORMATION:

APPLICANT: UHLEN, Mathias

TITLE OF INVENTION: SOLID PHASE DIAGNOSIS OF MEDICAL

CONDITIONS

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500

CITY: Alexandria

STATE: Virginia

COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA: US/08/477,270

APPLICATION NUMBER: US/08/477,270

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/261,010

FILING DATE:

APPLICATION NUMBER: US 07/781,157

FILING DATE: 07-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 16787/153 DFBC

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)836-9300

TELEFAX: (703)683-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 109 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-477-270-20

Query Match 2.38; Score 13; DB 1; Length 109;

Best Local Similarity 100.0%; Pred. No. 1.6e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 332 GADRSYCKPVEY 344

DB 30 GADRSYCKPVEY 42

Search completed: May 25, 2002, 22:25:58

Job time: 216 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2002, 22:24:02 ; Search time 21.52 seconds
(without alignments)
2482.605 Million cell updates/sec

Title: US-09-523-647-2
Perfect score: 556
Sequence: 1 MSKLRRVTVLALTSMAAC.....ILSSDTLSPVSDTENTHY 556

Scoring table: OLIGO Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size: 12

Total number of hits satisfying chosen parameters: 8

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: PIR 71.*
1: Pirl.*
2: Pirl.*
3: Pirl.*
4: Pirl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	556	100.0	556	1 S12602	60K cysteine-rich
2	556	100.0	556	2 A86560	60 kDa Cysteine-rich
3	45	8.3	557	1 B39439	60K cysteine-rich
4	45	8.1	558	2 JC5204	60K cysteine-rich
5	39	7.0	554	2 C81671	60 kDa outer membr
6	25	4.5	547	1 A32444	60K cysteine-rich
7	25	4.5	547	2 B43584	60K cysteine-rich
8	25	4.5	553	2 D71515	60K cysteine-rich

ALIGNMENTS

RESULT 1
S12602
60K cysteine-rich outer membrane protein precursor [similarity] - Chlamydia pneumoniae
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C:Date: 30-Sep-1993 #sequence revision 27-Jun-1994 #text_change 11-May-2000
C:Accession: S12602; H72063; A81604
R:Watson, M.W.; Al-Mandawi, S.; Lamden, P.R.; Clarke, I.N.
Nucleic Acids Res. 19, 5299, 1990
A:Title: The nucleotide sequence of the 60kDa cysteine rich outer membrane protein of Ch
A:Reference number: S12602; MUID:90384850
A:Accession: S12602
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-556 <WAT>
A:Cross-references: GB:X53511; NID:g550564; PIDN:CAA37590.1; PID:g550566
A:Experimental source: isolate IOL-207

R:Kelman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606
A:Accession: H72063
A:Molecule type: DNA
A:Residues: 1-556 <WAT>
A:Cross-references: GB:AE001640; GB:AE001363; MID:g4376845; PIDN:AAJ18697.1; PID:g4377
A:Experimental source: strain CWL029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heldelberg, J.F.; White, O.; Hicke
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A:Reference number: A81500; MUID:20150255
A:Accession: A81604
A:Molecule type: DNA
A:Residues: 1-556 <WAT>
A:Cross-references: GB:AE002180; GB:AE002161; NID:g7189117; PIDN:AAF38068.1; PID:g718
A:Experimental source: strain AR39, HL cells
C:Comment: This protein is associated with the differentiation of reticulate bodies i
It may also be an important virulence factor.
C:Genetics:
A:Gene: omcB; CP0195
C:Superfamily: 60K cysteine-rich outer membrane protein
C:Keywords: membrane protein; virulence
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-40/Domain: propeptide #status predicted <PRO>
F:41-556/Product: 60K cysteine-rich outer membrane protein #status predicted <WAT>

Query Match 100.0% Score 556; DB 1; Length 556;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKLRRVTVLALTSMAACFASGGIEAAVSAESLITKIVASAEKTPAPVPMATAKVVRLVR 60
Db 1 MSKLRRVTVLALTSMAACFASGGIEAAVSAESLITKIVASAEKTPAPVPMATAKVVRLVR 60
Qy 61 RNKQPVQKSRGAFCDKEFPCEEGRCQPVFAQQECSYGRLYSVKVNDDCNVEICQSVPE 120
Db 61 RNKQPVQKSRGAFCDKEFPCEEGRCQPVFAQQECSYGRLYSVKVNDDCNVEICQSVPE 120
Qy 121 YATVGSYPYPIELAIIGKDKCDVDVITQQLPCEAEFVSSDPETTPSDGKLVWKIDRLGAG 180
Db 121 YATVGSYPYPIELAIIGKDKCDVDVITQQLPCEAEFVSSDPETTPSDGKLVWKIDRLGAG 180
Qy 181 DKCKITVWKPLKEGCGCTAATVCACPELRSYTKCGOPATCIKQEGPDCACLRCPVCYKI 240
Db 181 DKCKITVWKPLKEGCGCTAATVCACPELRSYTKCGOPATCIKQEGPDCACLRCPVCYKI 240
Qy 241 EVNTGSAIARNVTVNDVPDGYSHASGORVLSNGLDMRPGDKKVTVEFFCQRGQIT 300
Db 241 EVNTGSAIARNVTVNDVPDGYSHASGORVLSNGLDMRPGDKKVTVEFFCQRGQIT 300
Qy 301 NVATVTCGGHKCSANVTTVVNEPCVQVNI SGADMSYVCKPVEYSISVSNPGDLVLHDVV 360
Db 301 NVATVTCGGHKCSANVTTVVNEPCVQVNI SGADMSYVCKPVEYSISVSNPGDLVLHDVV 360
Qy 361 IQDTLPSGVTVLEAPGGEICCNKVVWRKEMCPGETTQFLKLVKAQVPGRTTQVAVTSE 420
Db 361 IQDTLPSGVTVLEAPGGEICCNKVVWRKEMCPGETTQFLKLVKAQVPGRTTQVAVTSE 420
Qy 421 SNCGTCTSCAETTHHKGLAATHMGLDNDPICVGENTVYRICVTRNGSAEDTNNVSLIL 480
Db 421 SNCGTCTSCAETTHHKGLAATHMGLDNDPICVGENTVYRICVTRNGSAEDTNNVSLIL 480
Qy 481 KFSKELQPIASSGPTKGTISGNTVVDALPKLGSKESVEFSVTUKGIAPGDARGEAILSS 540
Db 481 KFSKELQPIASSGPTKGTISGNTVVDALPKLGSKESVEFSVTUKGIAPGDARGEAILSS 540
Qy 541 DTLTSPVSDTENTHY 556
Db 541 DTLTSPVSDTENTHY 556

RESULT 2

A86560
60 kDa Cysteine-rich omp [imported] - Chlamydomophila pneumoniae (strain J138)
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: A86560
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349
A:Accession: A86560
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-556 <STO>
A:Cross-references: GB:BA000008; NID:98978928; PIDN:BAA98763.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: omcB
C:Superfamily: 60K cysteine-rich outer membrane protein

Query Match 100.0%; Score 556; DB 2; Length 556;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSKLIIRVVTVLALTSMAFCFASGGIEAAVAESLITKIVASAEKTPAPVPMPTAKKRVLR 60
Db 1 MSKLIIRVVTVLALTSMAFCFASGGIEAAVAESLITKIVASAEKTPAPVPMPTAKKRVLR 60
QY 61 RNKQPVQKSRGAFCDKEFYPCCEGRQCPVEAQOESCGRLYSVKVNDCNVEICQSVPE 120
Db 61 RNKQPVQKSRGAFCDKEFYPCCEGRQCPVEAQOESCGRLYSVKVNDCNVEICQSVPE 120
QY 121 YATVGSYPPIEILAIKCKDCVDTITQQLPCEAEFVSSDPETPTSDGKLWKIDRLGAG 180
Db 121 YATVGSYPPIEILAIKCKDCVDTITQQLPCEAEFVSSDPETPTSDGKLWKIDRLGAG 180
QY 181 DKCKITVWVKPLKGGCCFTAAATVCACPELRSYTKCGQPAICIKOEGP 240
Db 181 DKCKITVWVKPLKGGCCFTAAATVCACPELRSYTKCGQPAICIKOEGP 240
QY 241 EVWATGSAIARNVTVDNVPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCFQRRGQIT 300
Db 241 EVWATGSAIARNVTVDNVPDGYSHASGORVLSFNLGDMRPGDKKVFTEFCFQRRGQIT 300
QY 301 NVATVTYCGGHKCSANVTTVNNEPCVQVNI SGADWSYVCKPVEYSISVSNPGDLVLHDVV 360
Db 301 NVATVTYCGGHKCSANVTTVNNEPCVQVNI SGADWSYVCKPVEYSISVSNPGDLVLHDVV 360
QY 361 IODTLPVSGTVLEAPGECCKVWRKEMCPGETLQFLVYKAQVPGRTNOAVTSE 420
Db 361 IODTLPVSGTVLEAPGECCKVWRKEMCPGETLQFLVYKAQVPGRTNOAVTSE 420
QY 421 SNGCTCTCAETTHWKGLAATHMCLDNDPFCVGTNTYRICVTRNGSAEDTNYSLIL 480
Db 421 SNGCTCTCAETTHWKGLAATHMCLDNDPFCVGTNTYRICVTRNGSAEDTNYSLIL 480
QY 481 KFSKELQPIASSGPTKTSGNTVDFDALPKLGSKEVSFVTLKGIAPGDARGEAILSS 540
Db 481 KFSKELQPIASSGPTKTSGNTVDFDALPKLGSKEVSFVTLKGIAPGDARGEAILSS 540
QY 541 DTLTSPVSDTENTHY 556
Db 541 DTLTSPVSDTENTHY 556

RESULT 3

B39439
60K cysteine-rich outer membrane protein 1 precursor - Chlamydomophila psittaci
C:Species: Chlamydomophila psittaci, Chlamydia psittaci
C>Date: 21-Feb-1992 #sequence_revision 27-Jun-1994 #text_change 31-Mar-2000

C:Accession: B39439; S12603
R:Everett, K.D.E.; Hatch, T.P.
J. Bacteriol. 173, 3821-3830, 1991
A:Title: Sequence analysis and lipid modification of the cysteine-rich envelope prot.
A:Reference number: A39439; MUID:91267949
A:Accession: B39439
A:Molecule type: DNA
A:Residues: 1-557 <EVS>
A:Cross-references: GB:M61116; NID:9144489; PIDN:AAB61619.1; PID:9144491
A:Experimental source: strain 68C
R:Watson, M.W.; Lambden, P.R.; Clarke, I.N.
Nucleic Acids Res. 18, 5300, 1990
A:Title: The nucleotide sequence of the 60kDa cysteine rich outer membrane protein of
A:Reference number: S12603; MUID:90384851
A:Accession: S12603
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-44, 'A', 46-72, 'E', 74-557 <WAT>
A:Cross-references: GB:X53512; NID:940625; PIDN:CAA37592.1; PID:940627
A:Experimental source: strain EAE/A22/M
C:Function:
O:Description: associated with the differentiation of reticulate bodies into elements
O:be an important virulence factor.
C:Superfamily: 60K cysteine-rich outer membrane protein
C:Keywords: membrane protein; virulence
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-40/Domain: propeptide #status predicted <PRO>
F:41-557/Product: 60K cysteine-rich outer membrane protein 1 #status predicted <MAT>
Query Match 8.3%; Score 46; DB 1; Length 557;
Best Local Similarity 100.0%; Pred. No. 2,7e-39;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 182 KCKITVWVKPLKGGCCFTAAATVCACPELRSYTKCGQPAICIKOEGP 227
Db 183 KCKITVWVKPLKGGCCFTAAATVCACPELRSYTKCGQPAICIKOEGP 228
RESULT 4
JC5204
60K cysteine-rich outer membrane protein 2 precursor - Chlamydomophila psittaci
C:Species: Chlamydomophila psittaci, Chlamydia psittaci
C>Date: 20-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 31-Mar-2000
C:Accession: JC5204
R:Hsia, R.; Bavoil, P.M.
Gene 176, 155-162, 1996
A:Title: Sequence analysis of the omp2 region of Chlamydia psittaci strain GPIC: Stru
A:Reference number: JC5203; MUID:97075924
A:Accession: JC5204
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-558 <HSI>
A:Cross-references: GB:U41759; NID:91783376; PIDN:AAB41143.1; PID:91783382
A:Experimental source: strain GPIC
C:Genetics:
A:Gene: omp2
C:Superfamily: 60K cysteine-rich outer membrane protein
C:Keywords: membrane protein; virulence
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-40/Domain: propeptide #status predicted <PRO>
F:41-558/Product: 60K cysteine-rich outer membrane protein 2 #status predicted <MAT>
Query Match 8.1%; Score 45; DB 2; Length 558;
Best Local Similarity 100.0%; Pred. No. 3e-38; Indels 0; Gaps 0;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 183 CKITVWVKPLKGGCCFTAAATVCACPELRSYTKCGQPAICIKOEGP 227
Db 185 CKITVWVKPLKGGCCFTAAATVCACPELRSYTKCGQPAICIKOEGP 229

R:Clarke, I.N.; Ward, M.E.; Lambden, P.R.
Gene 71, 307-314, 1988
A:Title: Molecular cloning and sequence analysis of a developmentally regulated cyste
A:Reference number: JT0419; MUID:89138006
A:Accession: JT0419
A:Molecule type: DNA
A:Residues: 30-547 <CIA>
A:Cross-references: GB:M35148; NID:g144485; PIDN:AAA23119.1; PID:g144487
A:Experimental source: serotype LI
C:Genetics:
A:Gene: omp2; omcB
C:Function:
A:Description: associated with differentiation of reticulate bodies into elementary b
A:Note: essential for the structural integrity of the outer envelope of the elementar
C:Superfamily: 60K cysteine-rich outer membrane protein
C:Keywords: membrane protein; virulence
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-40/Domain: propeptide #status experimental <PRO>
F:41-547/Product: 60K cysteine-rich outer membrane protein 2 #status experimental <MA

Query Match 4.5%; Score 25; DB 1; Length 547;
Best Local Similarity 100.0%; Pred. No. 2.3e-17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 KITVWVKPLKEGCGCFTAAATVCACPE 208
Db 175 KITVWVKPLKEGCGCFTAAATVCACPE 199

RESULT 7
B43584
60K cysteine-rich outer membrane protein 2 precursor, serotype E - Chlamydia trachoma
C:Species: Chlamydia trachomatis
C:Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 20-Aug-1999
C:Accession: B43584; S13120; S18979; S24275
R:de la Maza, L.M.; Fielder, T.J.; Carlson, E.J.; Markoff, B.A.; Peterson, E.M.
Infect. Immun. 59, 1196-1201, 1991
A:Title: Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodal
A:Reference number: A43584; MUID:91147205
A:Accession: B43584
A:Molecule type: DNA
A:Residues: 1-547
A:Cross-references: EMBL:X54389; NID:g40763; PIDN:CAA38259.1; PID:g40764
A:Experimental source: strain Bour, serotype E
R:Coles, A.M.; Allan, I.; Pearce, J.H.
Nucleic Acids Res. 18, 6713, 1990
A:Title: The nucleotide and derived amino acid sequence of the omp2 gene of Chlamydia
A:Reference number: S13120; MUID:91067486

A:Accession: S13120
A:Molecule type: DNA
A:Residues: 1-32, 'FT', 35-120, 'L', 122-131, 'A', 133-457, 'S', 459-547 <COL>
A:Cross-references: EMBL:X55903; NID:g40724; PIDN:CAA39396.1; PID:g40725
A:Experimental source: strain DK20, serotype E
C:Genetics:
A:Gene: omp2; omcB
C:Function:
A:Description: associated with differentiation of reticulate bodies into elementary b
A:Note: essential for the structural integrity of the outer envelope of the elementar
C:Superfamily: 60K cysteine-rich outer membrane protein
C:Keywords: membrane protein; virulence
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-40/Domain: propeptide #status predicted <PRO>
F:41-547/Product: 60K cysteine-rich outer membrane protein 2 #status predicted <MA>

Query Match 4.5%; Score 25; DB 2; Length 547;
Best Local Similarity 100.0%; Pred. No. 2.3e-17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 KITVWVKPLKEGCGCFTAAATVCACPE 208
Db 175 KITVWVKPLKEGCGCFTAAATVCACPE 199

RESULT 5
C81671
60 kDa outer membrane protein TC0727 [imported] - Chlamydia muridarum (strain Nigg)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000
C:Accession: C81671
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255
A:Accession: C81671
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-554 <TET>
A:Cross-references: GB:AE002341; GB:AE002160; NID:g7190754; PIDN:AAF39537.1; PID:g719075
A:Experimental source: strain Nigg (MoPn)
C:Genetics:
A:Gene: TC0727
C:Superfamily: 60K cysteine-rich outer membrane protein

Query Match 7.0%; Score 39; DB 2; Length 554;
Best Local Similarity 100.0%; Pred. No. 5.5e-32;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 451 DPICVGVTVYRICVTVNRGSAEDTVSLILKFSKELQPI 489
Db 449 DPICVGVTVYRICVTVNRGSAEDTVSLILKFSKELQPI 487

RESULT 6
A32244
60K cysteine-rich outer membrane protein 2 precursor, serotype LI and L2 - Chlamydia tra
C:Species: Chlamydia trachomatis
C:Date: 12-Oct-1989 #sequence_revision 27-Jun-1994 #text_change 16-Jul-1999
C:Accession: A32244; A43584; A36043; A30472; JT0419; S18981; S24277
R:Allen, J.E.; Stephens, R.S.
J. Bacteriol. 171, 285-291, 1989
A:Title: Identification by sequence analysis of two-site posttranslational processing of
A:Reference number: A32244; MUID:89123030
A:Accession: A32244
A:Molecule type: DNA
A:Residues: 1-547 <ALL>
A:Cross-references: GB:M23001; NID:g144552; PIDN:AAA23152.1; PID:g144553
A:Experimental source: strain L2/434/Bu
A:Note: parts of this sequence, including the amino ends of the precursor and mature pro
R:de la Maza, L.M.; Fielder, T.J.; Carlson, E.J.; Markoff, B.A.; Peterson, E.M.
Infect. Immun. 59, 1196-1201, 1991
A:Title: Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodalton
A:Reference number: A43584; MUID:91147205
A:Accession: A43584
A:Molecule type: DNA
A:Residues: 1-547
A:Cross-references: GB:M23001; NID:g144552; PIDN:AAA23152.1; PID:g144553
A:Experimental source: serovar 2, strain L2/434/Bu
R:Wahlberg, J.; Lundeborg, J.; Hultman, T.; Uhlen, M.
Proc. Natl. Acad. Sci. U.S.A. 87, 6569-6573, 1990
A:Title: General colorimetric method for DNA diagnostics allowing direct solid-phase gen
A:Reference number: A36043; MUID:90370827
A:Accession: A36043
A:Molecule type: DNA
A:Residues: 294-402 <WAH>
A:Experimental source: serotype L2
R:Lambden, P.R.; Everson, J.S.; Ward, M.E.; Clarke, I.N.
Gene 87, 105-112, 1990
A:Title: Sulfur-rich proteins of Chlamydia trachomatis: developmentally regulated transcr
A:Reference number: JQ0514; MUID:90236284
A:Accession: A30472
A:Molecule type: DNA
A:Residues: 1-46:528-547 <LAN>
A:Cross-references: GB:M35148; GB:M23180; GB:M35161; NID:g144485
A:Experimental source: serotype LI

Query Match 7.0%; Score 39; DB 2; Length 554;
Best Local Similarity 100.0%; Pred. No. 5.5e-32;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 451 DPICVGVTVYRICVTVNRGSAEDTVSLILKFSKELQPI 489
Db 449 DPICVGVTVYRICVTVNRGSAEDTVSLILKFSKELQPI 487

```

RESULT      8
D71515
60K cysteine-rich outer membrane protein 2 precursor serotypes B, C, and D - Chlamydia trachomatis
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 20-Aug-1999
C:Accession: D71515; C43584; S11673
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis
A:Reference number: A71570; MUID:59000809
A:Accession: D71515
A:Molecule type: DNA
A:Residues: 1-553 <ARN>
A:Cross-references: GB:AE001273; GB:AE001273; NID:g3328863; PIDN:AC68042.1; PID:g332887
A:Experimental source: serotype D, strain UW-3/Cx
R:de la Maza, L.M.; Fielder, T.J.; Carlson, E.J.; Markoff, B.A.; Peterson, E.M. Infect. Immun. 59, 1196-1201, 1991
A:Title: Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodalton protein
A:Reference number: A43584; MUID:91147205
A:Accession: C43584
A:Molecule type: DNA
A:Residues: 7-238 'V', 240-553 <DEL>
A:Cross-references: GB:X54388; NID:g40760; PIDN:CAA38257.1; PID:g40761
A:Experimental source: serotype C
R:Watson, M.W.; Lambden, P.R.; Ward, M.E.; Clarke, I.N. FEMS Microbiol. Lett. 65, 293-297, 1989
A:Title: Chlamydia trachomatis 60 kDa cysteine rich outer membrane protein: sequence homology
A:Reference number: S11673
A:Accession: S11673
A:Molecule type: DNA
A:Residues: 7-553 <WAT>
A:Cross-references: EMBL:X53510; NID:g40681; PIDN:CAA37588.1; PID:g40683
A:Experimental source: serotype B
C:Genetics:
A:Gene: omp2; omcB
C:Function:
A:Description: associated with differentiation of reticulate bodies into elementary bodies
A:Note: essential for the structural integrity of the outer envelope of the elementary body
C:Superfamily: 60K cysteine-rich outer membrane protein
C:Keywords: membrane protein; virulence
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-46/Domain: propeptide #status predicted <PRO>
F:47-553/Product: 60K cysteine-rich outer membrane protein 2 #status predicted <MAT>

```

```

Query Match      4.58; Score 25; DB 2; Length 553;
Best Local Similarity 100.0%; Pred. NO. 2.3e-17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 184 KITVWVKPLKGGCCFTAAATVCACPE 208
    |||||
DB 181 KITVWVKPLKGGCCFTAAATVCACPE 205
    |||||

```

```

Search completed: May 25, 2002, 22:26:30
JOB time: 148 sec

```

Tue May 28 08:55:15 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2002, 22:26:32 ; Search time 12.65 seconds
(without alignments)
1701.824 Million cell updates/sec

Title: US-09-523-647-2
Perfect score: 556
Sequence: 1 MSKLRRVTVALTSMASC.....ILSSDTLTSPVSDTENTHYV 556

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 12

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	556	100.0	556	1 OM6_CHLPN	P23700 chlamydia p
2	46	8.3	557	1 OM6_CHLPS	P23701 chlamydia p
3	25	4.5	547	1 OM6C_CHLTR	P26758 chlamydia t
4	25	4.5	547	1 OM6D_CHLTR	P18151 chlamydia t
5	25	4.5	547	1 OM6E_CHLTR	P23603 chlamydia t
6	25	4.5	547	1 OM6L_CHLTR	P21354 chlamydia t

ALIGNMENTS

RESULT 1	
ID OM6_CHLPN	STANDARD: PRT; 556 AA.
AC P23700: Q9TQI6;	
DT 01-NOV-1991 (Rel. 20, Created)	
DT 01-NOV-1991 (Rel. 20, Last annotation update)	
DE 16-OCT-2001 (Rel. 40, Last sequence update)	
DE 60 kDa outer membrane protein precursor (Cysteine-rich outer membrane protein) (CRP) (60 kDa cysteine-rich OMP).	
GN OM6C OR OMP2 OR CPN0557 OR CP0195.	
OS Chlamydia pneumoniae (Chlamydia pneumoniae).	
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.	
OX NCBI_TaxID=83558;	
RP SEQUENCE FROM N.A.	
RC STRAIN-IOL-207;	
RC MEDLINE=90384850; PubMed=2402463;	
RA Watson M.W., Al-Mahdawi S., Lamden P.R., Clarke I.N.;	
RT "The nucleotide sequence of the 60 kDa cysteine rich outer membrane protein of Chlamydia pneumoniae strain IOL-207."	
RL Nucleic Acids Res. 18:5299-5299(1990).	
RN [2]	

RP	SEQUENCE FROM N.A.
RC	STRAIN-CWL029;
RX	MEDLINE=99206606; PubMed=10192388;
RA	Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA	Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RA	"Comparative genomes of Chlamydia pneumoniae and C. trachomatis";
RL	Nat. Genet. 21:385-389(1999).
RN	[3]
RP	SEQUENCE FROM N.A.
RC	STRAIN=AR39;
RX	MEDLINE=20150255; PubMed=10684935;
RA	Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA	White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA	Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R., L.,
RA	Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA	Eisen J., Fraser C.M.;
RT	"Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT	pneumoniae AR39.";
RL	Nucleic Acids Res. 28:1397-1406(2000).
RN	[4]
RP	SEQUENCE FROM N.A.
RC	STRAIN=J138;
RX	MEDLINE=20330349; PubMed=10871362;
RA	Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA	Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RA	"Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT	from Japan and CWL029 from USA.";
RL	Nucleic Acids Res. 28:2311-2314(2000).
RN	[5]
RP	SEQUENCE FROM N.A.
RC	STRAIN=J138;
RA	Shirai M., Hirakawa H., Ouchi K., Tabuchi M., Kishi F., Kimoto M.,
RA	Takeuchi A., Nishida J., Shibata K., Fujinaga R., Yoneda H.,
RA	Matsushima H., Tanaka C., Furukawa S., Miura K., Nakazawa A.,
RA	Ishii K., Shiba T., Hattori M., Kuhara S.;
RT	"Comparison of outer membrane protein genes omp and pmp in the whole
RT	genome sequences of Chlamydia pneumoniae isolates from Japan and
RT	US.";
RL	Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC	-!- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES
CC	(RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL
CC	INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT
CC	VIRULENCE FACTOR.
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
EMBL; X53511; CAA37590.1; -	
EMBL; AEO01640; AAD18697.1; -	
EMBL; AEO02180; AAF38068.1; -	
EMBL; AP002547; BAA98763.1; -	
EMBL; AB033786; BAA85939.1; -	
PIR; S12602; S12602.	
PHCI-ZDPAGE; P23700; -	
TIGR; CP0195; -	
InterPro; IPR003506; Chlam_OMP6.	
PRINTS; PR01336; CHLAMIDIOM6.	
Outer membrane; Transmembrane; Signal; Virulence; Complete proteome.	
FT SIGNAL 1 22 POTENTIAL.	
FT PROPEP 23 40 POTENTIAL.	
FT CHAIN 41 556 60 KDA OUTER MEMBRANE PROTEIN.	
SEQUENCE 556 AA; 59719 MW; 8D7ED9234CC99458 CRC64;	
Query Match 100.0%; Score 556; DB 1; Length 556;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	

QY 1 MSLKLRVTVTLTSMASCFASGIEAAVESLITKIVASAETKPAVPMTAKKVLVR 60
DB 1 MSLKLRVTVTLTSMASCFASGIEAAVESLITKIVASAETKPAVPMTAKKVLVR 60
QY 61 RNKQPVQKSGARCDKFEFPCBGRQCPVEAQOQESYGRLYSVKVNDDCNVEICQSVPE 120
DB 61 RNKQPVQKSGARCDKFEFPCBGRQCPVEAQOQESYGRLYSVKVNDDCNVEICQSVPE 120
QY 121 YATVGSPIEILAIAGKKDCVDVITITQOLPCEAEFVSDPETTSKGKLVWKIDRLGAG 180
DB 121 YATVGSPIEILAIAGKKDCVDVITITQOLPCEAEFVSDPETTSKGKLVWKIDRLGAG 180
QY 181 DKCKITVWVPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKQEGPDCACLRCPVCYKI 240
DB 181 DKCKITVWVPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKQEGPDCACLRCPVCYKI 240
QY 241 EYVNTGSAIARNYTVDPVDPGYSHASGQVLSFNLGDMRPGDKVFTVEFCPQRRGQIT 300
DB 241 EYVNTGSAIARNYTVDPVDPGYSHASGQVLSFNLGDMRPGDKVFTVEFCPQRRGQIT 300
QY 301 NVATVYCGGHKCSANVTYVNEPCVOVNIAGADWSYVCKPVEYSISVSNPGDLVLHDVV 360
DB 301 NVATVYCGGHKCSANVTYVNEPCVOVNIAGADWSYVCKPVEYSISVSNPGDLVLHDVV 360
QY 361 IDOTLPSGVTVLEAPGGEICCNKVVNRIRKEMCPGETLQFKLVVKAQVGRFTNOVAVTSE 420
DB 361 IDOTLPSGVTVLEAPGGEICCNKVVNRIRKEMCPGETLQFKLVVKAQVGRFTNOVAVTSE 420
QY 421 SNGCTGCTCAETTHWKGAAATMCVLDNDPICVGNVTYRICVNTNRGSAEDTNVSLIL 480
DB 421 SNGCTGCTCAETTHWKGAAATMCVLDNDPICVGNVTYRICVNTNRGSAEDTNVSLIL 480
QY 481 KFSKELQPIASSGPTKGTISGNVTYVFDALPKLGSRESVEFSVTLKGIAPGDARGBAIILSS 540
DB 481 KFSKELQPIASSGPTKGTISGNVTYVFDALPKLGSRESVEFSVTLKGIAPGDARGBAIILSS 540
QY 541 DTLTSPVSDTENTHYV 556
DB 541 DTLTSPVSDTENTHYV 556

RESULT 2

OM6_CHLPS
ID OM6_CHLPS STANDARD; PRT; 557 AA.
AC P23701;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 60 kDa outer membrane protein precursor (Cysteine-rich outer membrane protein) (CRP) (60 kDa cysteine-rich OMP).
GN OMCB OR OMP2 OR ENVB.
OS Chlamydia psittaci (Chlamydia psittaci).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EA/A22/M;
RX MEDLINE=90384851; PubMed=2402464;
RA Watson M.W., Lambden P.R., Clarke I.N.;
RT "The nucleotide sequence of the 60 kDa cysteine rich outer membrane protein of Chlamydia psittaci strain EAE/A22/M.";
RL Nucleic Acids Res. 18:5300-5300(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=6BC;
RX MEDLINE=91267949; PubMed=2050637;
RA Everett K.D.E., Hatch T.P.;
RT "Sequence analysis and lipid modification of the cysteine-rich envelope proteins of Chlamydia psittaci 6BC.";
RL J. Bacteriol. 173:3821-3830(1991).
RN [3]
RP SEQUENCE FROM N.A.

RA Watson M.W.;
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT VIRULENCE FACTOR.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-slb.ch/announce/> or send an email to license@isb-slb.ch).
CC
CC EMBL: X53512; CAA37592.1; -;
DR EMBL: M61116; AAB61619.1; -;
DR PIR: S12603; S12603.
DR InterPro: IPR003506; ChlamOMP6.
DR PRINTS: PR01336; CHLAMIDIAOM6.
KW Outer membrane; Transmembrane; Signal; Virulence.
FT SIGNAL 1 22 POTENTIAL.
FT PROPEP 23 40 POTENTIAL.
FT CHAIN 41 557 60 KDA OUTER MEMBRANE PROTEIN.
FT VARIANT 45 45 A -> S (IN STRAIN 6BC).
FT VARIANT 73 73 E -> G (IN STRAIN 6BC).
SQ SEQUENCE 557 AA; 59843 MW; 0D444F09EAA073C6 CRC64;

Query Match

Best Local Similarity 8.3%; Score 46; DB 1; Length 557;
Matches 46; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 182 KCKITVWVPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKQEGP 227
DB 183 KCKITVWVPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKQEGP 228

RESULT 3

OM6C_CHLTR
ID OM6C_CHLTR STANDARD; PRT; 547 AA.
AC P26758;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 60 kDa outer membrane protein, serovar C precursor (Cysteine-rich outer membrane protein) (60-kDa CRP).
GN OMP2 OR OMP2B.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TW-3 / SEROVAR C;
RX MEDLINE=91147205; PubMed=1997423;
RA de la Maza L.M., Fiedler T.J., Carlson E.J., Markoff B.A., Peterson E.M.;
RT "Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodalton protein between the trachoma and lymphogranuloma venereum biovars of Chlamydia trachomatis.";
RL Infect. Immun. 59:1196-1201(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC Zhang Y.X., Caldwell H.D.;
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT VIRULENCE FACTOR.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X54388; CAA38257.1; -;
 DR EMBL; M85197; ABA23159.1; -;
 DR PIR; C43584; C43584;
 DR InterPro; IPR003506; Chlam_OMP6.
 DR PRINTS; PR01336; CHLAMIDIAOM6.
 KW Outer membrane; Transmembrane; Signal; Virulence.
 FT SIGNAL 1 22 POTENTIAL.
 FT PROPEP 23 40
 FT CHAIN 41 547 60 KDA OUTER MEMBRANE PROTEIN, SEROVAR C.
 FT SEQUENCE 547 AA; 58680 MW; 817BA5DC7FEA65D4 CRC64;
 SQ

Query Match 4.5%; Score 25; DB 1; Length 547;
 Best Local Similarity 100.0%; Pred. No. 1.1e-17;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 KITVWVKPLKEGCGCFTAAATVCACPE 208
 |||||
 DB 175 KITVWVKPLKEGCGCFTAAATVCACPE 199

RESULT 4
 OM6D_CHLTR STANDARD; PRT; 547 AA.
 ID OM6D_CHLTR
 AC P18151;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 60 kDa outer membrane protein precursor (Cysteine-rich outer membrane
 protein) (CRP) (60 kDa cysteine-rich OMP).
 DE OMCB OR OMP2 OR OMP2B OR CT443.
 GN Chlamydia trachomatis.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B/JAL120/OT;
 RX MEDLINE=90128208; PubMed=2612891;
 RA Watson M.W., Lambden P.R., Ward M.E., Clarke I.N.;
 RT "Chlamydia trachomatis 60 kDa cysteine rich outer membrane protein:
 sequence homology between trachoma and LGV biovars.";
 RL FEMS Microbiol. Lett. 53:293-297(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B/TW-05/OT;
 RX MEDLINE=91141306; PubMed=2287277;
 RA Allen J.E., Cerrone M.C., Beatty P.R., Stephens R.S.;
 RT "Cysteine-rich outer membrane proteins of Chlamydia trachomatis
 display compensatory sequence changes between biovariants.";
 RL Mol. Microbiol. 4:1543-1550(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D/UW-3/CX;
 RX MEDLINE=9900809; PubMed=9784136;
 RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
 RA Davis R.W.;
 RT "Genome sequence of an obligate intracellular pathogen of humans:
 Chlamydia trachomatis.";
 RL Science 282:754-759(1998).
 CC -!- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES
 CC (RBS) INTO ELEMENTARY BODIES (EBS). IT IS NECESSARY FOR STRUCTURAL
 CC INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT
 CC VIRULENCE FACTOR.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X53510; CAA37588.1; -;
 DR EMBL; AE001317; AAC68042.1; ALT_INIT.
 DR PIR; S11673; S11673.
 DR InterPro; IPR003506; Chlam_OMP6.
 DR PRINTS; PR01336; CHLAMIDIAOM6.
 KW Outer membrane; Transmembrane; Signal; Virulence; Complete proteome.
 FT SIGNAL 1 22 POTENTIAL.
 FT PROPEP 23 40
 FT CHAIN 41 547 60 KDA OUTER MEMBRANE PROTEIN.
 FT VARIAT 233 233 I -> V (IN STRAIN B/TW-05/OT).
 FT SEQUENCE 547 AA; 58694 MW; 42719B4BDCEDCEA CRC64;
 SQ

Query Match 4.5%; Score 25; DB 1; Length 547;
 Best Local Similarity 100.0%; Pred. No. 1.1e-17;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 KITVWVKPLKEGCGCFTAAATVCACPE 208
 |||||
 DB 175 KITVWVKPLKEGCGCFTAAATVCACPE 199

RESULT 5
 OM6E_CHLTR STANDARD; PRT; 547 AA.
 ID OM6E_CHLTR
 AC P23603;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 60 kDa outer membrane protein, serovar E precursor (Cysteine-rich
 outer membrane protein) (60-kDa CRP).
 GN Chlamydia trachomatis.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DK 20 / SEROVAR E;
 RX MEDLINE=91067486; PubMed=2251143;
 RA Colles A.M., Allan I., Pearce J.H.;
 RT "The nucleotide and derived amino acid sequence of the omp2 gene of
 Chlamydia trachomatis serovar E.";
 RL Nucleic Acids Res. 18:6713-6713(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BOUR / SEROVAR E;
 RX MEDLINE=91147205; PubMed=1997423;
 RA de la Maza L.M., Fiedler T.J., Carlson E.J., Markoff B.A.,
 RA Peterson E.M.;
 RT "Sequence diversity of the 60-kilodalton protein and of a putative
 15-kilodalton protein between the trachoma and lymphogranuloma
 venereum biovars of Chlamydia trachomatis.";
 RL Infect. Immun. 59:1196-1201(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC Zhang Y.X., Caldwell H.D.;
 RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES
 CC (RBS) INTO ELEMENTARY BODIES (EBS). IT IS NECESSARY FOR STRUCTURAL
 CC INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT
 CC VIRULENCE FACTOR.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
 CC -!- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF STRAIN DK 20.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; X55903; CAA39396.1; -;
 DR EMBL; X54389; CAA38259.1; -;
 DR EMBL; M85196; AAA23154.1; -;
 DR PIR; S13120; S13120.
 DR InterPro: IPR003506; Chlam_Omp6.
 DR PRINTS; PR01336; CHLAMIDIAOM6.
 KW Outer membrane; Transmembrane; Signal; Virulence.
 FT SIGNAL 1 22 POTENTIAL.
 FT PROPEP 23 40
 FT CHAIN 41 547 60 KDA OUTER MEMBRANE PROTEIN, SEROVAR E.
 FT VARIANT 33 34 SL -> FT (IN STRAIN DK20).
 FT VARIANT 121 121 I -> L (IN STRAIN DK20).
 FT VARIANT 132 132 V -> A (IN STRAIN DK20).
 FT VARIANT 458 458 N -> S (IN STRAIN DK20).
 SQ SEQUENCE 547 AA; 58708 MW; 0520656084F4E20AB CRC64;

Query Match 4.5%; Score 25; DB 1; Length 547;
 Best Local Similarity 100.0%; Pred. No. 1.le-17;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 KITVWVKPLKEGCGCFTAAATVCACPE 208
 DB 175 KITVWVKPLKEGCGCFTAAATVCACPE 199

RESULT 6
 OMSL_CHLTR STANDARD; PRT; 547 AA.
 AC P21354; P18586;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 60 kda outer membrane protein, serovars L1/L2/L3 precursor (Cysteine-
 DE rich outer membrane protein) (60-kda CRP).
 GN OMP2 OR OMP2B.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN-L2/434/BU; PubMed-2914847;
 RX MEDLINE-89123030; PubMed-2914847;
 RA Allen J.E., Stephens R.S.;
 RT "Identification by sequence analysis of two-site posttranslational
 RT processing of the cysteine-rich outer membrane protein 2 of Chlamydia
 RT trachomatis serovar L2.";
 RL J. Bacteriol. 171:285-291(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-404 / SEROVAR L3;
 RX MEDLINE-91147205; PubMed-1997423;
 RA de la Maza L.M., Fiedler T.J., Carlson E.J., Markoff B.A.,
 RA Peterson E.M.;
 RT "Sequence diversity of the 60-kilodalton protein and of a putative
 RT 15-kilodalton protein between the trachoma and lymphogranuloma
 RT venereum biovars of Chlamydia trachomatis.";
 RL Infect. Immun. 59:1196-1201(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-L1/440/LN;
 RX MEDLINE-89138006; PubMed-3066701;
 RA Clarke I.N., Ward M.E., Lambden P.R.;
 RT "Molecular cloning and sequence analysis of a developmentally
 RT regulated cysteine-rich outer membrane protein from Chlamydia
 RT trachomatis.";

RL Gene 71:307-314(1988).
 CC -|- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES
 CC (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL
 CC INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT
 CC VIRULENCE FACTOR.
 CC -|- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; M23001; AAA23152.1; -;
 DR EMBL; X54390; CAA38261.1; -;
 DR EMBL; M35148; AAA23119.1; ALT_INIT.
 DR PIR; A32244; A32244.
 DR PIR; S18981; S18981.
 DR PIR; A43584; A43584.
 DR PIR; JTO419; JTO419.
 DR Siena-2DPAGE; P21354; -;
 DR InterPro: IPR003506; Chlam_Omp6.
 DR PRINTS; PR01336; CHLAMIDIAOM6.
 KW Outer membrane; Transmembrane; Signal; Virulence.
 FT SIGNAL 1 22 POTENTIAL.
 FT PROPEP 23 40
 FT CHAIN 41 547 60 KDA OUTER MEMBRANE PROTEIN, SEROVARS
 FT L1/L2/L3.
 SQ SEQUENCE 547 AA; 58782 MW; 78CEB41CCF98472D CRC64;

Query Match 4.5%; Score 25; DB 1; Length 547;
 Best Local Similarity 100.0%; Pred. No. 1.le-17;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 KITVWVKPLKEGCGCFTAAATVCACPE 208
 DB 175 KITVWVKPLKEGCGCFTAAATVCACPE 199

Search completed: May 25, 2002, 22:30:06
 Job time: 214 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2002, 16:42:13 ; Search time 4990.55 Seconds
(without alignments)
8851.914 Million cell updates/sec

Title: US-09-523-647-1
Perfect score: 2111
Sequence: 1 tttagcaggtagtaggaga.....catcaaatctacatccgta 2111

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: gb_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vl.*
- 30: em_hgt_hum.*
- 31: em_hgt_inv.*
- 32: em_hgt_other.*
- 33: em_hgt_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description

	1	2111	100.0	2711	1	CPCRMP
	2	2107.8	99.8	11938	1	AE002180
C	3	2107.8	99.8	300550	1	AP002547
C	4	2106.2	99.8	10422	1	AE001640
	5	1671	79.2	1671	6	AX349525
	6	1185.2	56.1	3650	1	CHTCRPNV
	7	1185	56.1	1185	1	AF111201
	8	1183.6	56.1	2478	1	CPSCROMP
	9	1158.8	54.9	8950	1	CPU41759
	10	1156.6	54.8	3186	1	AF240773
	11	1031.8	48.9	1604	1	CPU76760
	12	944.8	44.8	1604	1	CPU76761
	13	889.6	42.1	2390	1	CHTOMPE
	14	888	42.1	2385	1	CTCROMP
	15	888	42.1	2390	1	CHTOMPA
C	16	888	42.1	11141	1	AE001317
	17	887.6	42.0	1734	1	AF304326
	18	887.6	42.0	1738	1	AF304327
	19	887.6	42.0	1752	1	AF304331
	20	887.2	42.0	1753	1	AF304328
	21	884	41.9	1753	1	AF304330
	22	882.8	41.8	1919	1	CHTOMPA2A
	23	881.4	41.8	1753	1	AF304329
	24	877.6	41.6	2345	1	CTSCCRP
	25	876.8	41.5	3012	1	CHTCRPA
C	26	876.6	41.5	12442	1	AE002341
	27	876.4	41.5	1726	1	CTOMP2
	28	876	41.5	2342	1	CTSECRP
	29	864.8	41.0	2430	1	CTSL3CRP
	30	817	38.7	1632	1	AF304332
	31	795.8	37.7	1185	1	AF111200
C	32	767.6	36.4	1947	6	AX338361
	33	725.8	34.4	1185	1	AF111199
C	34	719.6	34.1	1278	6	AX338362
	35	549.4	26.0	551	1	CPU56926
	36	544.6	25.8	551	1	CPU56925
	37	523.2	24.8	528	1	AF347609
	38	379.6	18.0	386	1	AF102831
	39	363.6	17.2	731	6	AX338251
	40	287.8	13.6	548	1	AF367407
	41	284	13.5	534	1	CPU56924
	42	282.4	13.4	534	1	CPU56927
	43	231	10.9	471	1	AF367405
	44	231	10.9	471	1	AF367406
	45	228.4	10.8	230	6	ARI44060

ALIGNMENTS

RESULT	1	CPCRMP	2711 bp	DNA	linear	BCT 06-NOV-1996
LOCUS		Chlamydia pneumoniae gene for cysteine rich outer membrane protein.				
DEFINITION		X53511				
ACCESSION		X53511.1				
VERSION		GI:550564				
KEYWORDS		outer membrane protein.				
SOURCE		Chlamydia pneumoniae.				
ORGANISM		Chlamydia pneumoniae				
REFERENCE		Bacteria: Chlamydiales; Chlamydiaceae; Chlamydia				
AUTHORS		Watson, M.W.				
TITLE		Direct Submission				
JOURNAL		Submitted (22-JUN-1990) Watson M.W., Southampton University England, Department of Microbiology, Faculty of Medicine, South Lab and Path block, Southampton General Hospital, Tremona road, Southampton, SO9 4XJ, UK				
REMARK		revised by [4]				
REFERENCE		2 (bases 1 to 1800)				
AUTHORS		Watson, M.W., al-Mahdawi, S., Lamden, P.R. and Clarke, I.N.				
TITLE		The nucleotide sequence of the 60 kDa cysteine rich outer membrane protein of Chlamydia pneumoniae strain IOL-207				
JOURNAL		Nucleic Acids Res. 18 (17), 5299 (1990)				

<div>90384850</div> <div>3 (bases 1 to 2711)</div> <div>Watson,M.W.</div> <div>Direct Submission</div> <div>Submitted (09-JAN-1991) M.W.Watson, Univ., Department of Microbiology, Southampton General Hospital, Southampton, S09 4XY, UK</div> <div>revised by [4]</div> <div>4 (bases 1 to 2711)</div> <div>Watson,M.W.</div> <div>Direct Submission</div> <div>Submitted (05-FEB-1993) M.W.Watson, Univ., Department of Microbiology, Southampton General Hospital, Southampton, S09 4XY, UK</div> <div>5 (bases 1 to 2711)</div> <div>Watson,M.W., Clarke,I.N., Everson,J.S. and Lambden,P.R.</div> <div>The Crp operon of Chlamydia psittaci and Chlamydia pneumoniae</div> <div>Microbiology 141 (pt 10), 2489-2497 (1995)</div> <div>96036206</div> <div>Location/Qualifiers</div> <div>1..2711</div> <div>/organism="Chlamyidophila pneumoniae"</div> <div>/strain="IOL-207"</div> <div>/db_xref="taxon:83558"</div> <div>prim_transcript 204..2571</div> <div>/gene="60kDa Crp"</div> <div>prim_transcript 204..643</div> <div>/gene="9kDa Crp"</div> <div>gene 204..645</div> <div>/gene="9kDa Crp"</div> <div>gene 204..2644</div> <div>/gene="60kDa Crp"</div> <div>prim_transcript 204..2509</div> <div>/gene="60kDa Crp"</div> <div>prim_transcript 204..2507</div> <div>/gene="60kDa Crp"</div> <div>prim_transcript 204..2508</div> <div>/gene="60kDa Crp"</div> <div>prim_transcript 204..644</div> <div>/gene="9kDa Crp"</div> <div>prim_transcript 204..645</div> <div>/gene="9kDa Crp"</div> <div>prim_transcript 204..2572</div> <div>/gene="60kDa Crp"</div> <div>prim_transcript 204..2573</div> <div>/gene="60kDa Crp"</div> <div>prim_transcript 204..2574</div> <div>/gene="60kDa Crp"</div> <div>prim_transcript 204..2575</div> <div>/gene="60kDa Crp"</div> <div>prim_transcript 204..2641</div> <div>/gene="60kDa Crp"</div> <div>prim_transcript 204..2642</div> <div>/gene="60kDa Crp"</div> <div>prim_transcript 204..2643</div> <div>/gene="60kDa Crp"</div> <div>prim_transcript 204..2644</div> <div>/gene="60kDa Crp"</div> <div>prim_transcript 204..2570</div> <div>/gene="60kDa Crp"</div> <div>291..297</div> <div>/gene="9kDa Crp"</div> <div>300..572</div> <div>/gene="9kDa Crp"</div> <div>/codon_start=1</div> <div>/transl_table=11</div> <div>/protein_id="CAA37589.1"</div> <div>/db_xref="GI:550565"</div> <div>/db_xref="SPTREMBL:Q46183"</div> <div>/translation="MKKAVLIAMFCGVVSLSSSRIVDCCFEDPCAPSSCNPCEVIR</div> <div>728..732</div> <div>/gene="60kDa Crp"</div> <div>RBS</div>		<div>739..2409</div> <div>/gene="60kDa Crp"</div> <div>/codon_start=1</div> <div>/transl_table=11</div> <div>/protein_id="CAA37590.1"</div> <div>/db_xref="GI:550566"</div> <div>/db_xref="SWISS-PROT:P23700"</div> <div>/translation="MSKLRRVVTVLALTSMAFCFSGGIEAAVAESLITKIIVASAEY</div> <div>-KPAPVMTAKKVLRRNPQVEKSGAFCDKEFYPCBEGRCOPVEAQOECYGRLY</div> <div>SYKVNDDCNVEICQSPYATVGSPIETILAIKKDCVDVVITQOLPCEAFVPSDD</div> <div>ETTPITSDGKLVMKIDRLGADKCIITVWKPLKGGCCFTAAITVCACPELRSYTKCCGP</div> <div>ATCIKQEGPDCACLRPCVCKIEVNTVGSIAIRNVTVDNPVPGYSHASGQVLSNL</div> <div>GMWRPKKVFTEPCPORGOITNAVTVTYCGGKHKCSANVTTVNPECPVOVNISAD</div> <div>WSYVCKPVEYSISVSNPGLDLVHDVQITLPSGVTVLEAPGGEICCNKVVWRKICD</div> <div>PGETQFKULVRAQVPGRETNAQVNTSENCGTCTSCAETTHMKGLAATHMCVLDTN</div> <div>DPICQVNTVYRICVNTNGSAQNTVLSLILKFSKELOPIASSGPTKGTITSGNTVVVFD</div> <div>LPKLGSKESVEFVTLKGIAPGDARGEAILSSDITLTPSVSDTENTHVV"</div> <div>BASE COUNT 824 a 502 c 546 g 839 t</div> <div>ORIGIN</div> <div>Query Match 100.0%; Score 2111; DB 1; Length 2711;</div> <div>Best Local Similarity 100.0%; Pred. No. 0;</div> <div>Matches 2111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;</div> <div>QY 1 ttgatacaggtagttaggagatgaataatctcctgactacatttaattcagataataaaccc 60</div> <div>DB 601 TTGATCAGGTAGTTAGGAGATGAATTAATTCCTGACTACCTTAATTCAGATAATAAACCC 660</div> <div>QY 61 aaatgttgagggttaagagtttacaaaacattctaccatgagcagagagaaaaataaac 120</div> <div>DB 661 AAATGTTGAGGTTAAGAGTTTACAAAACATTCATCCCATGGCAGAGAAAAATAAAC 720</div> <div>QY 121 atgcgataaggatcccttatgtccaaactcatcagacagtagttacgttccttgcgcta 180</div> <div>DB 721 ATGCGATAGGAGATCCCTATGTCCAACTCATCAGCAGTAGTTACGGTCTCTCGGCTA 780</div> <div>QY 181 acgagtagcgagttgcttgcagcgggggtatagggcgctgtgagagatctctg 240</div> <div>DB 781 ACGAGTAGGCGAGTTGCTTTGCCAGCGGGGTATAGAGGCGCTGTAGCAGAGTCTCTG 840</div> <div>QY 241 attactaagatcgtcgctagtcgcgaaacaaagcagcacctgttccatatacagcagag 300</div> <div>DB 841 ATTACTAAGATCGTCGCTAGTGGGAAACAAAGCCAGCACCTGTCTTATCAGACCGAAG 900</div> <div>QY 301 aaggttagactgtccgttagaataaaacacagttgaacaaaaagcgttggtgctttt 360</div> <div>DB 901 AAGGTTAGACTTGTCCGTAGAAATAAACACACAGTTGAACAAAAAGCCGTGGTGCTTTT 960</div> <div>QY 361 tgtgataaagaattttatccctgtgaagaggagcagatgcaacctgtagaggtcagcaa 420</div> <div>DB 961 TGTGATAAAGAAATTTTATCCCTGTGAAGAGGAGCATGTCAACCTGTAGAGGCTCAGCAA 1020</div> <div>QY 421 gactctgtacggaagattgtattctgtaaaagtaaacgatgattgcaacgtagaatt 480</div> <div>DB 1021 GAGTCTTGCTACGAGAGATTGTATCTGTAAAGTAAACAGATGATTGCAAGGTAGAAAT 1080</div> <div>QY 481 tgcagtccttccagaatacgtactgttaggattcctctaccctattgaaatccttgcct 540</div> <div>DB 1081 TGCCAGTCCGTTCCAGAAATACGCTACTGTAGGATCTCTTACCCCTATTGAATCCTTGC 1140</div> <div>QY 541 ataggcaaaaaagattgt 600</div> <div>DB 1141 ATAGGCAAAAAAGATTGT 1200</div> <div>QY 601 ttctgaagcagtagatccagaaacaaactcctacaagtgtatgggaaattagtcgtgaaatc 660</div> <div>DB 1201 TTCGTAAAGCAGTATCCAGAAACAACTCCCTACAGTGTAGGAAATAGTCTGCAAAATC 1260</div> <div>QY 661 gatcgctgggtgcaggagataaatgcaaaattactgtatgggtgaaacctcttaagaa 720</div> <div>DB 1261 GATCGCCTGGGTCCAGGAGATAAATGCAAAATTTACTGTATGGGTAAACCTCTTAAGAA 1320</div>
--	--	---

QY 721 ggttgcttcacagctgctactgtatgtgcttgcgcagagctcccggtttattactataa 780
|||||
Db 1321 ggttgcttcacagctgctactgtatgtgcttgcgcagagctcccggtttattactataa 1380
|||||
QY 781 tgggtcaaccagcatttatttaagaagaagacactgactgtgctgctaagatgc 840
|||||
Db 1381 tggcggtcaaccagcatttatttaagaagaagacactgactgtgctgctaagatgc 1440
|||||
QY 841 cctgtatgctacaaatcgaaagtagtgaacacagagatgctgattggcccgtaacgttaact 900
|||||
Db 1441 cctgtatgctacaaatcgaaagtagtgaacacagagatgctgattggcccgtaacgttaact 1500
|||||
QY 901 gtatgaatacctgttcccgatggctattctcatgcatctgggtcaagagttctctcttt 960
|||||
Db 1501 gtatgaatacctgttcccgatggctattctcatgcatctgggtcaagagttctctcttt 1560
|||||
QY 961 aacttgaagagacatgagacctggcgataaaaaggattttacacagttgagttcgcctcaaa 1020
|||||
Db 1561 aacttgaagagacatgagacctggcgataaaaaggattttacacagttgagttcgcctcaaa 1620
|||||
QY 1021 agaagaggtcaaatcactaaactgctgactgtactgtactgtactgtactgtactgtactgt 1080
|||||
Db 1621 agaagaggtcaaatcactaaactgctgactgtactgtactgtactgtactgtactgtactgt 1680
|||||
QY 1081 gcaaatgtaactacagttgttaagcctgtgtacaaagtaaatatctctgtgctgat 1140
|||||
Db 1681 gcaaatgtaactacagttgttaagcctgtgtacaaagtaaatatctctgtgctgat 1740
|||||
QY 1141 tgggttacgtatgtaaacctgtggagtaactctactctactctactctactctactctgagacttg 1200
|||||
Db 1741 tgggttacgtatgtaaacctgtggagtaactctactctactctactctactctactctgagacttg 1800
|||||
QY 1201 gttcttcagatgctgatacgaagatacactccctctctgtgttacagtaactcgaagct 1260
|||||
Db 1801 gttcttcagatgctgatacgaagatacactccctctctgtgttacagtaactcgaagct 1860
|||||
QY 1261 cctggtgagagatgctgttaataaagcttcttggtggtatttaagaagatgtgccagga 1320
|||||
Db 1861 cctggtgagagatgctgttaataaagcttcttggtggtatttaagaagatgtgccagga 1920
|||||
QY 1321 gaaacccctcaagttaaactgtagtgaagctcaagttccctctggaagattcacaaatacaa 1380
|||||
Db 1921 gaaacccctcaagttaaactgtagtgaagctcaagttccctctggaagattcacaaatacaa 1980
|||||
QY 1381 gttgcagtaactagtgagttcaactgcggaacatgatacttgcgcagaaacaacaaca 1440
|||||
Db 1981 gttgcagtaactagtgagttcaactgcggaacatgatacttgcgcagaaacaacaaca 2040
|||||
QY 1441 cattggaaggtcttcagctaccatctgctgattatgacacacaaatgactcctatctgt 1500
|||||
Db 2041 cattggaaggtcttcagctaccatctgctgattatgacacacaaatgactcctatctgt 2100
|||||
QY 1501 gtagggaataactgctctatogtatctgtgtaactaaacgttggttctcgaagatact 1560
|||||
Db 2101 gtagggaataactgctctatogtatctgtgtaactaaacgttggttctcgaagatact 2160
|||||
QY 1561 aacgtatcttaactctgaagttctcaaaagaacttcagccaatgcttctcgaagttcca 1620
|||||
Db 2161 aacgtatcttaactctgaagttctcaaaagaacttcagccaatgcttctcgaagttcca 2220
|||||
QY 1621 actaaaggaacatttcaggtataacccgttcttctcgaacgttttaactaaactcgttct 1680
|||||
Db 2221 actaaaggaacatttcaggtataacccgttcttctcgaacgttttaactaaactcgttct 2280
|||||
QY 1681 aagaatctgtagagttttctgttaccttgaaggtattgtccccggagatgctcgcgc 1740
|||||
Db 2281 aagaatctgtagagttttctgttaccttgaaggtattgtccccggagatgctcgcgc 2340
|||||
QY 1741 gaagctattcttctctgtacactgacttcacagtatcagacacagaaaaataccac 1800
|||||
Db 2341 gaagctattcttctctgtacactgacttcacagtatcagacacagaaaaataccac 2400
|||||
QY 1801 gttgattaaattcttaaggaattatcttaaacgagcagatattccgctctgtaggat 1860
|||||

Db 2401 GTGTATTAATTTCTAAGGAATTATCTCTAAGCAGACCGATATTCCTGCTCTGCTTTAGGAT 2460
|||||
QY 1861 agctttcaagaagtagtaccgcttttagtaccttaacgtactaaagcggttttttggtttata 1920
|||||
Db 2461 AGCTTTCAAAGAGTAGTACCGCTTTTAGTACCTTACGTACTAAAGCGGTTTTTTTCTTTATA 2520
|||||
QY 1921 agctcttcaatccaatcgtagagtttcttaatacaagataattttaagtttctgaaatc 1980
|||||
Db 2521 AGCTCTTCAATCCCAATCGTAGAGTTCTTAAATCAAGATATTTAATTTCTGAAATC 2580
|||||
QY 1981 ctaagattatttttaaaagcccatcttttaggtatgtaataaaaaattttaataaagct 2040
|||||
Db 2581 CTAAGATTTATTTAAAGCCCATCTTTTAGGTATGTAATTAATAATTTTAATTAAGCT 2640
|||||
QY 2041 ttctcagtgtaacctgcttcttaggaactacactaggaacggtatgcatcaaatc 2100
|||||
Db 2641 TTTCTAGTGTAACTGCTTCTTTTAGGAACCTACACTAGGAGACGGTATGTCATCAATC 2700
|||||
QY 2101 tacatcccgta 2111
|||||
Db 2701 TACATCCCGTA 2711
|||||
RESULT 2
AE002180 11938 bp DNA linear BCT 30-MAY-2000
LOCUS Chlamydomophila pneumoniae AR39, section 16 of 94 of the complete
DEFINITION genome
ACCESSION AE002180 AE002161
VERSION AE002180.2 GI:8163382
KEYWORDS
SOURCE Chlamydomophila pneumoniae AR39.
ORGANISM Bacteria: Chlamydiales; Chlamydiaceae; Chlamydomophila.
REFERENCE 1 (bases 1 to 11938)
AUTHORS Read,T.D., Brunham,R.C., Shen,C., Gill,S.R., Heidelberg,J.F.,
White,O., Hickey,E.K., Peterson,J., Umayam,L.A., Utterback,T.,
Berry,K., Bass,S., Linher,K., Weidman,J., Khouri,H., Craven,B.,
Bowman,C., Dodson,R., Gwinn,M., Nelson,W., DeBoy,R., Kolonay,J.,
McClarty,G., Salzberg,S.L., Eisen,J. and Fraser,C.M.
TITLE Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
pneumoniae AR39
JOURNAL Nucleic Acids Res. 28 (6), 1397-1406 (2000)
MEDLINE 20150255
PUBMED 10684935
REFERENCE 2 (bases 1 to 11938)
AUTHORS Read,T.D., Brunham,R.C., Shen,C., Gill,S.R., Heidelberg,J.F.,
White,O., Hickey,E.K., Peterson,J., Umayam,L.A., Utterback,T.,
Berry,K., Bass,S., Linher,K., Weidman,J., Khouri,H., Craven,B.,
Bowman,C., Dodson,R., Gwinn,M., Nelson,W., DeBoy,R., Kolonay,J.,
McClarty,G., Salzberg,S.L., Eisen,J. and Fraser,C.M.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
COMMENT On Jun 1, 2000 this sequence version replaced gi:7189117.
FEATURES Location/Qualifiers
source 1..11938
/organism="Chlamydomophila pneumoniae AR39"
/strain="AR39"
/db_xref="taxon:115711"
/note="synonym: Chlamydia pneumoniae AR39"
73..1839
/gene="CP0187"
73..1839
/gene="CP0187"
/note="similar to GB:M54884 SP:P21893 PID:147550
PID:887842 GB:U00096; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="single-stranded-DNA-specific exonuclease RecJ"
/protein_id="AAF38060.1"

```

/db_xref="GI:7189118"
/translation="MTNSDNAAGLLWAHPKEDPAFLCMIIEKHELPPTVAQIFISR
GFQTIOEHFLYSHLSLYDPLGLDMSKAVERLLLRKEHVMYIGSDVDGDMTG
VALLVEFRDIDVHVSYFELGAILKQHGETSLIAKKEEGITLLITVDCGITAGKEV
SDITRQGDVITIDHMPGKIPHCVALTNPKLRDHTYNNRELTGVGVAFLARGVLN
ALISNVPKSGSLKULDLVITGTTIDVGLLGENVMVRYGKEIARGARGLNK
ICALCGVEKSESTDIVLKAPKNSLGRDPAKGVLELLTODDERVDALIMELDN
INRQRLEAVFDQVQVILNSPEIKQAAIVLSSTAWHARVPIISARLAKYNKP
VVIATQRIKGKSGARTIGSPLLGVKCSLLLSYGGHFAAGVINKEDKVBDFKK
KFVHLVNSLKGDTLPHLEIDAVADDAIDYDILLASHELFEPFGKGLMPFIYSKVR
QVRYPKVLPGNHLKLYLSOKERNLEGVAFGLGRHADALKASWHYPLETAYTPRLSOTS
GSGVHLLVRDRISSEPRSD"
2093..3214
/gene="CP0188"
CDS
2093..3214
/gene="CP0188"
/translation="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF38061.1"
/db_xref="GI:7189119"
/translation="MSIAIAREQHAAILDMHPKPSIAFSESQARTSWEKRAHPYLY
RLEIITGVVFKLGLIFFTPLGLFWLOKICNFIILGAGGWIFRPICRDSNLLRQA
YAARLFSQDQHVSVRVCLQYDFVFDGLRLRPNAKPKRMVLSNNGSCLDEYR
TVLQGEKDWIFRAESQSNILFINIPGVKSGNITRNNVKSYQACVYRLRDEPAG
POARQIVAYGYSLGASQVAKSKEIADGSDSVRFVFKDGRGARTGAVAKOFIGSLG
VMLANTHNNISERKSOLHCPFLFYIGKDSQGNLIGDLFKKTCFAAPFLDPKNL
ECSGKGIPTAOTGLRHRHILSDDDVIKAVGHIRRHFN"
3725..4261
/gene="CP0189"
CDS
3725..4261
/gene="CP0189"
/translation="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF38062.1"
/db_xref="GI:7189120"
/translation="MACQHEGCVLEERIEIDIKSDTKWVSITQAAKLHNVTQRA
IYVAKOKKLKASKETREIDIKLEEKRNRYSRKSLYQCELVFDNGKCYSTNOV
AQIIGIPQVKYVYATRTGIRGERKGAAWIHVSEIERKYNEYLKQAAKLLKGAEPK
EHOAPNEPPEIPEFESN"
4517..6034
/gene="CP0190"
CDS
4517..6034
/gene="CP0190"
/translation="similar to PID:1783380; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="glutamyl-tRNA synthetase"
/protein_id="AAF38063.1"
/db_xref="GI:7189121"
/translation="MWNENVRVAVSPGDPHYGTAYMALFNEIFAKRFGKMLIRI
EDTDRTRSDVDEENFESALRWCGIORDGPDGPGYPGPRQSERTKIYGYVETLLK
TDCAYKCEATPOELAEARAVASTLGRGGVDRRYRVLSPPEVASREAAQGPYIRLKV
PLSGCEVFEDYSGKRVYFPWADVDQVLYKSDGFPYHFANVIDHLMGITHVLRGEF
NLSSTPKHLLLEAFGEWPEPFLHMLPNPDGTLKSRKNPTSIFYRYRDSGYVKEAF
VNFLLMGYSMDGEVYSLERIETFPNPRIGKCGAVFDIOKLDMNKHLYLNHEGSP
ECLLKEQGLMDEFFLKLPLQSRITTLAEFINLTSFSGLLLEYRVBELLPQAL
SPEKAAILLYSVYKYLEKTDQWTKETCYLGSKLQAQFNVRHKKAILPLLYVATTCGK
QGLPLFDSIEIGLKPRARALVYAEKLLGVGPKLAATVDRKMQREDFEATFDL"
complement(6199..6387)
/gene="CP0191"
CDS
complement(6199..6387)
/gene="CP0191"
/translation="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11

```

```

/product="conserved hypothetical protein"
/protein_id="AAF38064.1"
/db_xref="GI:7189122"
/translation="MKLFYFCNIIFLTSLFSVSCHTTTPPRSYILAQGRTPLVNKAHS
LEFFLARSVFNTCYNTNL"
6645..6752
/gene="CP0192"
CDS
6645..6752
/gene="CP0192"
/translation="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF38065.1"
/db_xref="GI:7189123"
/translation="MFSKKSFNKKIYLEKIIDKIFIKFICNIVHGNVLY"
6829..7101
/gene="CP0193"
CDS
6829..7101
/gene="CP0193"
/translation="similar to PID:1783381; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="cysteine rich outer membrane protein 3"
/protein_id="AAF38066.1"
/db_xref="GI:7189124"
/translation="MKKAVLIAAMFCGVVSLSCCRIVDCCFEDPCAPSSCNPCEVIR
KKERSCGGNACGSYVPCSNPCGSCNCSQSPQVKCTSPDGRCKQ"
complement(7165..7281)
/gene="CP0194"
CDS
complement(7165..7281)
/gene="CP0194"
/translation="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF38067.1"
/db_xref="GI:7189125"
/translation="MSLDIGISYRMFIFFCHRVCFVNSYPOHLGLLSELR"
7268..8938
/gene="CP0195"
CDS
7268..8938
/gene="CP0195"
/translation="similar to SP:P23700 PID:550566 PID:1326243
PID:1326245 GB:AB001363; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="60 kDa outer membrane protein"
/protein_id="AAF38068.1"
/db_xref="GI:7189126"
/translation="MSKLIIRVVTVLALTSMAFCASGGIEAAVAESLITKIVASAET
KPAPVMTAKRLVRNRKQPVQKSRGAFCEFYPCEEGRCPVPAQOQESCVGRLY
SVVNDNCNVEICQSVPEYATVGSPIEILAIKKDCVDVVIITQOLPCRAEYSSDP
ETPTTSDGLKVMRIDRLGADGDKITVWPKLEGCCFTAAATCAPELKSYTKCGOP
AICIKQEGPDCACLPCVCKIEVNTGSAIARNVDPNPVPCYSHASQGVLSGNL
GDMRPDKKVTVEFCPORRGOITNVATVYTCGGHKCSANVTTVVNBPCQVNVISGAL
WSYVCKPVEYSISVNPGLVLDHVVQDTLPDPSGVTVLEAPGGEICCNKVYWRKEMC
PGETLOFKLVKRAQVGRFTNQVAVTSESNGTCTSCAETTHWKGLAATHMCLVDTN
DPLCGENTVYRIVTRNGSAEDTNVSLIKFSKELOPIASSGPTKGTIGTSGNTVVFDA
LPLKGSKEVSEFVSVLKGIAPGDARGEAILSSDTLTSPVSDTENTHY"
9217..9807
/gene="CP0196"
CDS
9217..9807
/gene="CP0196"
/translation="similar to SP:P18587; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="15 kDa cysteine-rich outer membrane protein"

```

Tue May 28 08:55:07 2002

/protein_id="AAF38069.1"									
/db_xref="gi:7189127"									
/translation="MSSNLHPVGGTGTGAAPESVLNIVEIAASGSVTAGLQAITSS									
PGMYNLIGWAKTFPIQPIRESKFIQSRACQITLLVLGILLVWAGLACMFIFHSQIGA									
NAFLIIPAAIIGLLKLLVTSICFDEACTISKLWFOKWAGVLEQLDGLINNSKIF									
GHVKTGEGTSRAITPVLNDGRGTPLVSLPLSKIARV"									
complement(9885..11831)									
gene	99.8%	Score 2107.8	DB 1:	Length 11938					
Query Match	99.9%	Pred No. 0:							
Best Local Similarity	99.9%	Mismatches 2:	Indels 0:	Gaps 0:					
Matches 2109:	Conservative	0:							
QY	1	ttgatcaggtagtagagagatgaattaatccctgactacttaattcagataataaacc	60						
DB	7130	TTGNTCAGGTAGTAGGAGATGAATTAATCCCTGACTACCTTAATTCAGATATAAACCC	7189						
QY	61	aaatgttgagggaagatttacaacaattctaccgatggcagagaagaaaaataaac	120						
DB	7190	AAATGTTGAGGGTAAGAGATTACAAAACATCTACCCGATGGCAGAGAAAAATAAAC	7249						
QY	121	atcgatagagatccctatgtccaaactcatcagacagtagtagttagcgtccctgcgcta	180						
DB	7250	ATGCGATAGGAGATCCCTATGTCCAAACTCATCAGACGAGTAGTACGGTCCCTTGCCTA	7309						
QY	181	acgagtagcggagtagtcttggcagcggggtatagagcgcgtgtagcagagtcctctg	240						
DB	7310	ACGAGTATGGCGAGTGTCTTGGCCAGCGGGGTATAGAGGCCGTGTAGCAGAGTCTCTG	7369						
QY	241	attactaagatcgtcgtcagtagtgcgaaacaaagccagcacctgttccctatacacagcgaag	300						
DB	7370	ATTACTAAGATCGTCGTAGTCGCGAAACAAAGCCAGCACCTGTTCCTATGACACGCGAAG	7429						
QY	301	aaggttagactgtccgtagaaataaacaaccagttgaacaaaagcgtgggtctttt	360						
DB	7430	AAGGTTAGACTTGTCCGTAGAATAAACACAGTTGAACAAAAAGCCGTGCTCTTT	7489						
QY	361	tgtgataaagaattttatccctgtggaagaggagcagatgtcaaccctgtagagactcagcaa	420						
DB	7490	TGTGATTAAGAATTTTATCCCTGTGAAGAGGGACGATGTCAACCTGTAGAGGCTCAGCAA	7549						
QY	421	gagcttgctacgaagattgtattctgttaaaagtaaacgattgttgcgaacgtagaatt	480						
DB	7550	GAGTCTTGCTACGGAGATTTGTATTCTGTAAAGASTAAACGATGATTGCAACGTAGAANA	7609						
QY	481	tgccagtcctgtccagaatcacgtactgttaggatctcttaccctattgaaatcctgtct	540						
DB	7610	TGCCAGTCCGTTCCAGAAACGCTACTGTAGGATCTCTTACCTATTGAAATCCCTTGCT	7669						
QY	541	ataggcaaaaaagattgtgttgattgtgtattacacaacagctacaccttgcgaagctgaa	600						
DB	7670	ATAGCAAAAAAGATTGTGTGATTGTGTGATTACACACACAGCTACCTTGGGAAGCTGAA	7729						
QY	601	ttcgaacagctgacccagaacaactcctacaagtgatgggaattagtcgtgaaatac	660						
DB	7730	TTGCTAACAGCTGATCCAGAACAACTCCTACAAGTGTATGGGAAATTAGTCTGGAAATC	7789						
QY	661	gatcgctgggtgcaggagataaattgtaaaattactgtatgggtaaaaacctcttaagaa	720						
DB	7790	GATCGCTGGGTGTCAGGAGATAAATGCAAAATTTACTGTATGGGTAAACCTCTTAAAGAA	7849						
QY	721	ggttgccttcacagctgactgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	780						
DB	7850	GGTGTGCTGCTTACAGCTGCTACTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	7909						
QY	781	tcggttcaaccagccatttatttaagaagaagagacactgactgtgctgctgctgaagtc	840						
DB	7910	TGCGGTCAACGCCATTGTATTAAAGCAAGAGGACCTGACTGTGCTTGCCTTAAGATGC	7969						
QY	841	cctgtatgtacaaaatcaagtagtgaacacagagatctgtattgtcccgtaacgttaact	900						
DB	7970	CCTGTATGCTACAAAATCGAAGTAGTGAACACAGGATCTGCTATTGCCGTAACTAACT	8029						
QY	901	gtagataatcctgttcccgatggctattctctatgcatcgttggtgcaagagatctctctttt	960						
DB	8030	GTAGATAATCCTGTTCCCGATGGCTATCTCATGCATCTGTCCTCAAGAGTCTCTCTTTT	8089						
QY	961	aacttagagacatgagacacctggcgataaaaaggtatttacagttgattctgacctcaa	1020						
DB	8090	AACITTAGGACATGAGACCTTGGCGATAAAAAGGATATTACAGTTGAGTTGTGCGCCCAA	8149						
QY	1021	agaagaggtcaaatcaactaacgttgcctactgttaacttactcgtcgtggtgacacaaatgttct	1080						
DB	8150	AGAAGAGGTCAAACTACTTAACGTTTGTACTGTACTGTACTGTACTGTACTGTACTGTACT	8209						
QY	1081	gcaaatgtaactacacagttgtaaatgagccttgggtgtgtacaaagtaaatatctctgggtcgtat	1140						
DB	8210	GCAATGTAACTACAGTTGTTAATGAGCCTTGTGTACAAAGTAAATATCTCTGGTGTCTGAT	8269						
QY	1141	tggttctacgtatgtaaacctgtgagactctactcactcagtatcgtaacccctgggagacttg	1200						
DB	8270	TGCTCTTACGTATGTAAACCTGTGGAGTACTCTATCTCAGTATCGAATCCTTGAGACTGTG	8329						
QY	1201	gttctcatgtagtgcgtgacccaagatacacctccctctgtgtgttacagttactcgaagct	1260						
DB	8330	GTCTTTCATGATGTCGTGATCCAAAGATACACTCCCTTCTGGTGTACAGTACTCGAAGCT	8389						
QY	1261	cctggtggagagatcgcgtgtaataaagttgttggcgatatataaagaaatgtgcccagga	1320						
DB	8390	CCTGGTGGAGAGATCTGCTGTATAAAGTGTGTTGGCGTATTAAAGAAATGTGCCCAGGA	8449						
QY	1321	gaacccctccagtttaaacctgttagtgaagctcaagttcctctggtgttacagttactcgaatcaa	1380						
DB	8450	GAAACCTCCAGTTAACTTGTAGTGAAGCTCAAGTCTCTGGAAGATTTCACAAATCAA	8509						
QY	1381	gttcagtaactagtgagtgtaactgcgggaacatgtacatcttgcgcagaaaaacaaca	1440						
DB	8510	GTTGCAGTAACCTAGTGAAGTAACTGCGGAACATGTACATCTTGCSCAGAAAAACAACA	8569						
QY	1441	catggaagctcttgcagctacccatattgtgcgtattagacacaaatgatcctatcgt	1500						
DB	8570	CATTGGAAGGCTTGCAGCTACCCATATGTGCGTATTAGACACAAATGATCTATCTGT	8629						
QY	1501	gtaggagaaaaactactcgtctatcgatcgttaactaacctgggtctcgtcgaagataact	1560						
DB	8630	GTAGGAGAAAACTACTGTCTATCTGTAATCTCAAGTAATTCAGCCAAATAGCTTCTTCAGGTCCA	8749						
QY	1561	aacgtatcttaactctgaagttctcaaaagaaacttcagccaataagctcttcaggtcca	1620						
DB	8690	AACGTATCTTTAATCTTGAAGTCTCAAAAGAACTTCAGCCAAATAGCTTCTTCAGGTCCA	8749						
QY	1621	actaaagggaacgatttcagggttaataccgtgtgttgcagcgtttacctaataactcgtttct	1680						
DB	8750	ACTAAAGGAACGATTTTACGGTAATACCGTGTGTTTTCGACGCTTTACCTAAACTCGGTTCT	8809						
QY	1681	aaggaatctgtagagtttctgttaccttgaaggtattgtcccgagagatgctcgcgcgc	1740						
DB	8810	AAGGAATCTGATAGATTCTGTGTACCTTGAAGGTATTGCTCCCGAGATGCTCCGCGC	8869						
QY	1741	gaagctattcttctctctgatacaactgacttcaaccagttacacacacagaaaaataccac	1800						
DB	8870	GAAGCTATTCTTCTTCTGTATACACTGACTTCCAGTATCAGACACAGAAAAATACCCAC	8929						
QY	1801	gtgtataaattctaaagaaattatcctaaagcagagcgatattcccgctcgtctttaggat	1860						
DB	8930	GTGTATTAAATCTTAAGGAATATCTTAAAGCAGAGCGATATTCCGCTCTGCTTTAGGAT	8989						
QY	1861	agcttcaagaagtagcgccttagtaccttaactaactaaagcgggtttttttttttttata	1920						
DB	8990	AGCTTCAAGAACTACCGCTTTAGTACTTACGTACTTAAAGCGGTTTTTTTGTGTTTATA	9049						
QY	1921	agcttctcaatccaatgtaggtttcttaatacaagatatatttaagttcttcaaatc	1980						
DB	9050	AGCTCTTCAATCCAAATCGTAGAGTTCTTATCAAGAGATATTATTAAAGTTCTTGAATC	9109						
QY	1981	ctaagatttatatttaaaagcccatcttttaggtatgttaattaaatttttaataagct	2040						

```

Db 9110 CTAAGATTATTTTAAAGCCCATCTTTTATAGGTATGTAATTAATAATTTTAAATGCT 9169
Qy 2041 ttctcctagtgtaacctgtcttcttaggaactacactaggagaacggtatgtcatcaaatc 2100
|||||tttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
Db 9170 TTTCCCTAGTGAACCTGCTTCTTTAGGAACATACACTAGGAGAACGGTATGTCATCAAAATC 9229
Qy 2101 tacatcccgta 2111
|||||tttttttttt
Db 9230 TACATCCCGTA 9240

RESULT 3
AP002547/c
LOCUS
DEFINITION
Chlamydomophila pneumoniae J138 300550 bp DNA linear BCT 08-DEC-2000
section 3/4.
ACCESSION AP002547 AB0333786 AB0333787 AB0333816 AB0333817 BA0000008
VERSION AP002547.2 GI:10176693
KEYWORDS
SOURCE
Chlamydomophila pneumoniae J138 (strain:J138) DNA.
ORGANISM
Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.
REFERENCE
1 (sites)
Shirai,M., Hirakawa,H., Kimoto,M., Tabuchi,M., Tabuchi,M., Kishi,F., Ouchi,K.,
Shiba,T., Ishii,K., Hattori,M., Kuhara,S. and Nakazawa,T.
Comparison of whole genome sequences of Chlamydia pneumoniae J138
from Japan and CWL029 from USA
Nucleic Acids Res. 28 (12), 2311-2314 (2000)
20330349
2 (bases 1 to 300550)
Shirai,M.
Direct Submission
Submitted (04-JUL-2000) Mutsunori Shirai, Yamaguchi University
School of Medicine, Department of Microbiology; 1-1-1
Minamikuoshi, Ube, Yamaguchi 755-8505, Japan
(E-mail:meshirai@po.cc.yamaguchi-u.ac.jp, Tel:81-836-22-2227,
Fax:81-836-22-2415)
On Sep 15, 2000 this sequence version replaced gi:6172298
91:6172300 gi:6172396 gi:6172398 gi:8978869.
AB033786-AB033787, AB033816-AB033817: Submitted (25-Oct-1999).
FEATURES
source
1..300550
/organism="Chlamydomophila pneumoniae J138"
/strain="J138"
/db_xref="taxon:138677"
/note="synonym:Chlamydia pneumoniae (strain J138)"
complement(21..794)
/gene="dapP"
complement(21..794)
/gene="dapP"
/codon_start=1
/transl_table=11
/product="diaminopimelate epimerase"
/protein_id="BAA98725.1"
/db_xref="GI:8978890"
/translation="MAYFSESTISKYFIYSGAGNRFLGELTPEYEDVRFLLCQETRV
DGFLLKSSCADQLIFNSDGRPTMCGNLRCALAHLSQKGSIDSVDSDGLYS
GYFYSMDRLVMDTLADWRASVHRLESRPDLPKVCVCIHTGPHAVVILPEISL
SLGDFLRYHTSPDGVNVNVQILCHQLRVRYERGVETGTAACGTGALASALV
SNSYGWKESIOHTWGGELMTVSNRGRVYLGQSVTRDL"
complement(763..1338)
/gene="clpP.1"
complement(763..1338)
/gene="clpP.1"
/codon_start=1
/transl_table=11
/product="Clp protease"
/protein_id="BAA98726.1"
/db_xref="GI:8978891"
/translation="MADGEVHKLRDIIEKELLKARRVFESEPTKESADAIKKIWL
ELKDGKIVFVINSFGSVNDAFVMDQIKMLTSPVTVWVGLAASMGVSLCAAP
GRFATPSRIMIHOPSIGGPIFGQATDLDIHARELTAKRIIDVVEATNQPRDII
EKAIARDWMWTANEAKDGLLDGLIFSNL"

```

```

gene
CDS
complement(1358..2917)
/gene="glyA"
complement(1358..2917)
/gene="glyA"
/codon_start=1
/transl_table=11
/product="serine hydroxymethyltransferase"
/protein_id="BAA98727.1"
/db_xref="GI:8978892"
/translation="MLKVEPKPKFAIVEFTKVAVVSLLUKLENASGKKGOSLAS
TAYLAALDHLNLNAPSIGERLIDELKSORSHLKMIASBNYSSLSVLAMGNLLTDKYC
EGSPKRPYSCCENVDATMECVETLFAADCAQVPHSGCADANLLANVNLTKHY
QGPVASKYGTCTNEUTEEETTLKAEKSSCVCLGPSUNSGCHLTHGNVRUNLTKHY
RCFYDYNPOTCEFDYAEISRLAKYKPKVLIAGYSSYRRLNFAVLQKLAIEYESTLNG
VDMAHFAGLVAGVGVFDEENPIYADIIVTTTHKTLRGPRGGLVLATREYESTLNG
CPLMMGGPLPHVIAAKTVALKEALSVDFFKYAHQVNVNNAERLAEFSLHGLRLTGT
DNHMMVLDLGSLSGKIAEDILSSVGIANRNSLPSDAICGWDTSGTRLGTPALTTL
GMGIDEEVADIIIVKVLNRILRSLCHVEGSSKKNKGELPEATAOERDVRVRLRLRFP
LYPEIDLEALV"
3107..3778
/gene="CP10522"
3107..3778
/gene="CP10522"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAA98728.1"
/db_xref="GI:8978893"
/translation="MTLYLGLNQTKARKYOAHYLPILTLPYAKSTPONKRALOFLPO
ATHVLTSPSTHLSRMTSLSKATLTKYLCIGESTKERLLSLFGOVKYVYVATQ
BIAGKIFPLQAQPSARILYPHSSILARPVIREFLYNRFTFSPHYTVKPKLKNLI
LSYKKIKIIFTPSTVRAFAKIFPRFPKTYWCGRMTLQEFQKFSQKQVSLLETGK
SRTSP"
3843..4175
/gene="CP10523"
3843..4175
/gene="CP10523"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAA98729.1"
/db_xref="GI:8978894"
/translation="MASSATPGDGTAPSLPPATPRPRYNFKLALFVTIALVWIAL
IATTAIGLCTHPLCSPFILTALPYFISRYCSHYARNVYIALDVVDPHKLQDMKS
HSPFSDR"
4223..5302
/gene="CP10524"
4223..5302
/gene="CP10524"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAA98730.1"
/db_xref="GI:8978895"
/translation="MSGPSRTESSQSVLSYVPRKEIAPKQFTIAKILTALIASL
ALGALVAGISITIVGNPVELALLITLTFSVVTLVLYHOMTSKYSSNMQVLEQNF
PLGAMKEKNVDCYSNEMQYNNHLPKFKVAIQTDASQEQPTFLTLGLRVIEKNQST
GIIFNPVPTNLIDNTNLSTILYSTLAKDKSVMDTKQKQGGKAGDEPSPTEYRV
VKLNEALDQFNLSNLSAEKSKILPTFLGHVCGPKSELNQOQYRQALLAYENCL
KAATESHAIIVALPLFTSVYVEVPPEELLPKEGTFYWDNTOAFCKRALLDAIQNTALR
YQPSRLVILQDPENTIESORSEE"
5642..6406
/gene="CP10525"
5642..6406
/gene="CP10525"
/codon_start=1
/transl_table=11
/product="CT398 hypothetical protein"
/protein_id="BAA98731.1"
/db_xref="GI:8978896"
/translation="MHDALSLIAIOELDTMIRLMRVKKEHQELAKVQSLKSDIR
KVOEKEIEMENLKTQIRDGNRIQIESEQIKNLENOAAVKKMDEFNALTQEMTTANK
ERRSLRHQLSLDMDKQAGGEDLIVLSKESLASTENSSSVIEKEIFESIKKINEECKAL

```



```

gene
complement(3421..5091)
/gene="omcB"
/note="Cp0557"
complement(3421..5091)
/gene="omcB"
/codon_start=1
/transl_table=11
/product="60 kDa Cysteine-Rich OMP"
/protein_id="AADI8697.1"
/db_xref="GI:4376849"
/transl_table="MSKILRRVTVIALSMASCFAGGIEAAVSLTIKTVASAEAT
KPAVPMTAKKVLRRNKQPVQEKRGAFCKDEFFYPEEGRCQPVAAQOESCYGLR
SVKYNDDCNLVQSVPEYATVGSPPYIEILAIGKKCDVDVITQQLPEARAFVSDDP
ETTDPGKLVKIDRLAGADKCKITVWPKLGGCCFATAACVCELPKRSYTKCQOP
AICIKQSGPCACIKPVCYKIEVNTGSAIARNVTVDNVPDGYSHASQORVLNFNL
GDMRPGCKFTVEFCQRCQKITNNATVTCYGHKCSANVTTVNAEPCVQVNVISGL
WSYVCKPVEYSISVSNFGDLVLDHVDITQLPSGVTVLEAPGGEICCNKYVRKIDCM
PGETLQFLKLVKVAQVGRFNRQNAVTSSESNCGTSCAETTHWKLAAATHMCLVDTN
DPICVGVNTVYRICVTVNGSAEDTNVSLIKFKSKELQPIASSGPTKGTISGNTVWFEA
"

```

PGETLQFLVYRAQVFEKRLNG
 DPICVGTNYRIVCTNRGSAEDTNVLIKF
 LKSKELQFIASSGTTKGTISGNTVFV
 LKSGSEVFEVSTLKGAPDARGEAILSD
 TLTSPVSDTENTHY" complement(5258..5530)

```

gene
    /gene="omcA"
    /note="CPn0558"
    complement(5258..5530)
GDS

```

```

/codon_table=1
/transl_table=11
/product="9 kDa-Cysteine-Rich Lipoprotein"
/protein_id="AAD18698.1"

```

```

/protein_id=ACAD10030.1
/db_xref=GI:4376950
/translation="MKKAVLIAMFCGVVLSSSCRIVDCGFEDPCAPSSCNPCVEIF
KKRSCGGNACGSYPVSCNPGCTECNSPQVKGCTSPDGRCK"
5972 6160

```

```

gene
  /gene="Cp00559"
CDS
  /gene="Cp00559"
  5972..6160

```

```
/codon_start=1  
/transl_table=11  
/product="CT44.1 hypothetical protein"  
/protein_id="AMD18699.1"
```

```

gene
/translation="MKLFYFCNIIFLTSLFSVCSCTHTPPRSYILAQGRTPLVNKAH
LEFFLARSVENTCYNTNL"
complement(6325..7842)
/db_xref="GI:4376851"

```

CDS

```

/codon_start=1
/transl_table=11
/product="Glutamyl-tRNA synthetase"

```

```
/db_xref="GI:4376852"
/translation="MNNENVRVVAPSPGDPVGTAYMALFNEIPAKRFKGMILF
EDTTRSRQDVEENIFSAALRWCIQWDEGPDVGPPYQSRSTKIYQGVETLL
TTCAYKCFATPOBELAEMAVASTLGYRGYDRRYLSPVEVASREAAGQPTIRLF
```

PUSGECFEDI SAGUW F PWDVDDYDZT
 WLSSTPKHLLLYAFGWEPPEV LHMPLDLLNPDGT KLSKRKNPTS IFYRDSHYVKE
 VNFELTCMYSMEGDVEVSLERII IETFNRRIGKSAVDIQKLDMWNKHYYLNHEGG
 VNFELTCMYSMEGDVEVSLERII IETFNRRIGKSAVDIQKLDMWNKHYYLNHEGG

gene
complement(8098. .8634)
QG1PFDSEILGKPRARLRVYAEKLLGVPKLAATVDFEMQREDEEAIFDL
SPKAAIILYSYVKYLEKTOWTRETICILGSNWLQAQFNVHNRGRREAFDL

CDS

```

/cocon_start=1
/transl_table=11
.
/product="CHLPS Euo Protein"
--id="AA018701.1"

```

```
/db_xref="GI:4376853"
/translation="MACEQHEGCELEEREIEEDIKDSITKWSITQAAKLHNVTFR
```

Db 4569 GATCGCTGGTGCAGGAGATAAATCAAAATTAATGTATGGTAAACCTCTTTAAAGAA 4510
Qy 721 gttgtgcttcaagctgctactgtatgctgtgcccagagctcggttcttactataa 780
Db 4509 GGTGTGCTGCTTCACAGCTGCTACTGTATGCTGTGCTGCCAGAGCTCCGTCTTATACATAA 4450
Qy 781 tgcgggtcaaccagccattgttataagcaagaagagacctgactgctgtgctcctaagatgc 840
Db 4449 TGGCGGTCAACGACCATTTGTATTAAAGCAAGAAGGACCTGACTGTGCTTGCCTAAGATGC 4390
Qy 841 cctgtatgctcaaaaactcaagctagtgaaacacagatgctgtattgcccgttaacgttaact 900
Db 4389 CCTGTATGCTTACAAATTCGAGTAGTAACACAGGATCTGCTATTGGCCGTAACGTAACT 4330
Qy 901 gtatagataactcctgtctcccgatggctattctcatgcatctggtgtcgaagagttctctcttt 960
Db 4329 GTAGATAAATCCTGTTCCGATGGCTATTCTCATGCTATCTGGTCAAAAGAGTTCTCTCTTT 4270
Qy 961 aacttaggagacatgagacctggcgataaaaaggLatlacagttgagttctgcccctcaa 1020
Db 4269 AACTTAGGACATGAGACCTGGCGATAAAAAGGTATTACAGTTCACTTCTGCCCTCAA 4210
Qy 1021 agaagaggtcaaatcaactgaacttactactgaacttactcgtggtgacacaaatgttct 1080
Db 4209 AGAAGAGGTCAATCACTAACGTTGCTACTGTACTGTACTTCTGCGTGGACACAAATGTCT 4150
Qy 1081 gcaaatgtaactacagttgttaatatgagcctgtgttacagtaaatatctctgtgctgat 1140
Db 4149 GCAAAATGAATCACTACGTTGTTAATGAGCCTTGTGTACAAGTAATATCTCTGGTGCCTGAT 4090
Qy 1141 tggcttcaagctatgataaacctgtggagctactctactcagatcgaatccctggagacttg 1200
Db 4089 TGGTCTTACGTATGAACCTGTGGAGTACTCTATCTCAGTATCGAATCTCTGGAGACTTG 4030
Qy 1201 gtctctcatatgctgtatcaagatacaactccctctgtgtgttacaagtactcgaagct 1260
Db 4029 GTTCTCTCATGATGCTGTGATCCAGATACACTCCCTCTCTGCTGTACAGTACTCGAAGCT 3970
Qy 1261 cctggtggagagatcctgctgtataaaagtgtgttgctgtattaaagaatgtgccaggga 1320
Db 3969 CTTGGTGGAGAGATCTGCTGTATAAAGTGTGTGGCGCTATTAAAGAAATGTGCCAGGA 3910
Qy 1321 gaaacctccagtttaaacctgttagtgaagctcaagttccctgggaagattcacaaaacaa 1380
Db 3909 GAAACCTCCAGTTTAAACTGTGTAGTGAAGCTCAAGTTCCTCGGAAGATTACAAAACAA 3850
Qy 1381 gtgcagtaactagtgagttcaactcgggaacatgtacattcttgcgagaaacaaacaa 1440
Db 3849 GTTGCAGTAATAGTGAGTCTAACTCGGAACATGTATCATCTTGGCGAGAAACAAACAA 3790
Qy 1441 cattggaaaaggtcttcagctaccatattgtcgtattagacacaaaatgtacctatctgt 1500
Db 3789 CATTTGAAAAGGTCTTGCAGCTACCCATATGTGCGTATTAGACACAAATGTATCTGT 3730
Qy 1501 gtaggagaaaactgtctatcgtactgttaactaaacgtggttctgctggaagatact 1560
Db 3729 GTAGGAAATAACTGTCTATCGTATCTGTATCACTAAACCGTGTCTGTCTGAAGTACT 3670
Qy 1561 aacgtactcttaactcttgagttctcaaaagaacttcagcaaatagcttcttcagggtcca 1620
Db 3669 AACGTATCTTTAATCTTGAAAGTTCTCAAAAGAACTTCAGCCAAATAGCTTCTTCAGGTCCA 3610
Qy 1621 actaaaggaacgattccaggtataacacgtgtttctgcagcgtttacacctaaactcgttct 1680
Db 3609 ACTAAAGAAACGATTTTCAGGTAAATACCGTGTGTTTTCGACGCTTTTACCTTAAACTCGGTCT 3550
Qy 1681 aaggaactgtagagtttctgttaccttgaaggtattgtctcccgagatgctcgcggc 1740
Db 3549 AAGGAATCTAGAGTTTCTGTACTCTGAAGGTATTGCTCCCGGAGATGCTCGCGGC 3490
Qy 1741 gaagctattcttctctgtatcacactgacttaccagttatcagacacagaaaataccccc 1800
Db 3489 GAAGCTATTCTTCTTCTGATACACTGACTTCCACGATTACACAGATATCAGACAGAAAAATCCGAC 3430

Qy 1801 gtgatttaaatcttaagaataattatactaaagcagagcgatattccgctctgctcttaggat 1860
Db 3429 GTGTATTAATTTCTTAAGGAATATCTCTAAAGCAGAGCGATATTCGCGTCTGCTTTAGGAT 3370
Qy 1861 agctttcaagaaglacgcgttttagtaccttacactaaagcgggtttttttttttttata 1920
Db 3369 AGCTTTCAAGAAGTACCGCTTTAGTACCTTACGTACTAAAGCGGTTTTTTCTTTTATA 3310
Qy 1921 agctcttcaatccaatcagttcttaatacaagaatattatttaagtttctcgaatc 1980
Db 3309 AGCTCTTCAATCCAATCTAGAGTTCTTAATCAAAATATTTAATGTTTCTGAAATC 3250
Qy 1981 ctaagatttatttttaaaagccatcttttttaggtatgataataaaatcttcttaataagct 2040
Db 3249 CTAAGATTATTTTAAAAAGCCCATCTTTTAAAGTATGTAATTAATAATTAATGT 3190
Qy 2041 ttctctagtgtaacctgtcttttaggaactcacactaggagaaacgggtatgctcatcaaatc 2100
Db 3189 TTTCTCTAGTGTAACTGCTCTCTTTAGGAACCTACACTAGGAGACGGTATGTCATCAATC 3130
Qy 2101 tacatcccgta 2111
Db 3129 TACATCCCGTA 3119
RESULT 5
AX349525
LOCUS
DEFINITION Sequence 48 from Patent WO0202606.
ACCESSION AX349525
VERSION AX349525.1 GI:18615369
KEYWORDS
SOURCE Chlamydomophila pneumoniae.
ORGANISM Chlamydomophila pneumoniae.
REFERENCE 1 (sites)
AUTHORS Ratti,G. and Grandi,G.
TITLE Immunisation against Chlamydia pneumoniae
JOURNAL Patent: WO 0202606-A 48 10-JAN-2002;
Chiron S.p.A. (IT)
FEATURES
Location/Qualifiers
source 1. 1671
/organism="Chlamydomophila pneumoniae"
/db_xref="taxon:83558"
BASE COUNT 488 a 342 c 370 g 471 t
ORIGIN
Query Match 79.2%; Score 1671; DB 6; Length 1671;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1671; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 139 atgtccaaactcatcagcagagtagttacggtcccttgcgttaacgagatcgcgagttgc 198
Db 1 ATGTCCAAACTCATCAGACGAGTAGTTACGCTCTTGGCTAACGAGTATGCGAGTTGC 60
Qy 199 ttgcccagcgggggtatagagggccgctgtagcagagctctctgtactactaagatcgct 258
Db 61 TTGTCCAGCGGGGTATAGAGCCCGCTAGCAGAGTCTCTGTATTAAGATTCGTCGCT 120
Qy 259 agtgcggaacaaacccagcagcctgttctctatgacagcgaaggttagactgtccgt 318
Db 121 AGTGGGAAACAAAGCCAGCACCTGTCTTATGACAGCGAAGAGGTAGACTTGTCCGT 180
Qy 319 agaaataaacaccagtttgaacaaaagcgtggtgctttctgtgtataaagaattttat 378
Db 181 AGAAATAAACACCACTGTGAACAAAAAGCCGTGCTGCTTTTGTGATAAGAAATTTAT 240
Qy 379 cccgtggaaggggagcagtgatgcaacctgtagaggtcagcaagagttcttctcaggaaga 438
Db 241 CCTGTGAAGAGGGAGCATGTCACCTGTAGAGGCTCAGCAAGAGTCTTCTCTACGGAAGA 300

QY	439	ttgtattctgtataaagtaaacgtagtgcgaacgtatagaaatttgcaggtccgttcacagaa	498	Db	1381	TATCGTATCTGTAACTAACCGTGGTTCGTGCTGAAGATATAACGATATCTTTAACTTGG	1440
Db	301	TTGTATTCTGTAAAAGTAAACGATGATTGCAACGTAGAAATTTGCCAGTCCGTTCCAGAA	360	QY	1579	aagttctcaaaagaacttcagcaaatagctcttcaggtcccaactaaagaagatttca	1638
QY	499	tacgtactgttaggatctccctaccctattgaaatcccttgctataggaacaaagattgt	558	Db	1441	AAAGTTCTCAAAAGAACTTCAGCCAAATAGTCTTCAGGTCCCACTAAAGGAACGATTTC	1500
Db	361	TACGTACTGTAGGATCCTTACCTATTGAAATCCTTGCTATATAGCAAAAAGATGT	420	QY	1639	gataatacgtgtgttttcgacgtttacactaaactcggtcttaaggaaatctgtagaattt	1698
QY	559	gttgatgttgattacacacacagctactctgcgaagctgaattcgtaaacagtgatcca	618	Db	1501	GGTAATACCGTTGTTTTCGACGCTTTACTAACTTCGTTCTTAAGGAATCTGTAGAGTTT	1560
Db	421	GTTGATGTTGTATTACACAACAGCTACCTTCGGAAGCTGAATTCGTAAAGCAGTGATCCA	480	QY	1699	tctgttaccttgaagggtattgtctcccgagatgctcgcgcgaagctattcttctctt	1758
QY	619	gaacaactcctacaagtgatggaaattagctctggaataatcgatcgctgggtgcagga	678	Db	1561	TCTGTTACCTTGAAGAGTATTGCTCCCGAGATGCTCCGGCGAAGCTATTCTTCTTCT	1620
Db	481	GAACAACCTCTACAAGTGATGGAAATTAGTCTGGAAATCGATCGCTGGGTGCAGGA	540	QY	1759	gataactgactccaccagtatcacacacagaaataaccacgtgtattaa	1809
QY	679	gataaatgcaaaattactgtatgggtataaacctcttaagaagggtgtgcttccacagt	738	Db	1621	GATACACTGACTTCACCACTATCAGACACAGAAAATACCACGCTGTATTAA	1671
Db	541	GATAATGCAAAATTACTGTATGGGTAAACCTCTTAAAGAAGGTTGCTGCTTCACAGCT	600	RESULT	6		
QY	739	gctactgtatgtctgtccagagctccgtctcttataactaaatgcggtcaaccagccatt	798	CHTCRPNV		CHTCRPNV	
Db	601	GCTACTGTATGTGCTGCCAGAGCTCCGTTCTTATCTTACTTAATGCGGTCAACACGCCAT	660	LOCUS		Chlamydia psittaci 6BC low molecular weight cysteine-rich envelope	
QY	799	tgtattaagcaagaagacgtgactgtgctgtcctaagatgcctgtatgtacaaaatc	858	DEFINITION		protein (envA), high molecular weight cysteine-rich envelope	
Db	661	TGTATTAAAGCAAGACCTGACTGTGCTTGCTTAAGATGCCCTGTATGCTACAAAATC	720	ACCESSION		M61116	
QY	859	gaagtagtgaacacagatctgctattgccgaagtaactagatagatccctgttccc	918	VERSION		M61116.1	GI:144489
Db	721	GAAGTAGTGAACACAGATCTGCTATTGCCCCGTAACTACTGTAGATAATCTGTTCC	780	KEYWORDS		high molecular weight cysteine-rich envelope protein; sulfur-rich protein.	
QY	919	gatgctattctcatgactcgttgcgaagagttctctttaaacttagagacatgaga	978	SOURCE		Chlamydia psittaci	
Db	781	GATGCTATTCTCATGCATCTGGTCAAGAGTCTCTCTTTTAACTTAGGAGACATGAGA	840	ORGANISM		Bacteria; Chlamydiales; Chlamydiaceae; Chlamydiaophila.	
QY	979	cctggcgataaaaaggtattacagttgagttcgcctcaagaagaggtcacaatcaat	1038	REFERENCE		1 (bases 1 to 3650)	
Db	841	CCTGCGATATAAAGGTATTTACAGTTGAGTCTGCGCTCAAGAAGAGGTCAATCACT	900	AUTHORS		Everett, K.D.E. and Hatch, T.P.	
QY	1039	aacgtgtactgtactactcgtcgttggtggacacaaatttctgcataatgttaactacagt	1098	TITLE		Sequence analysis and lipid modification of the cysteine-rich envelope proteins of Chlamydia psittaci 6BC	
Db	901	AACGTGCTACTGTAACTTACTGCGGTGGACACAAATTTCTGCAAAATGTAACTACAGTT	960	JOURNAL		J. Bacteriol. 173, 3821-3830 (1991)	
QY	1099	gttaatgagcctgtgtacaaagaataatctctgctgctgattgctctacgtatgataa	1158	MEDLINE		91267949	
Db	961	GTTAATGAGCCTTGTGTACAAAGTAATAATCTCTGCTGTGTTGCTTACGTATGTAAA	1020	FEATURES		Location/Qualifiers	
QY	1159	ctgtgagtaactctatctcagtatcgaaatcctgcgagacttggttctctatgactgtg	1218	1..3650		/organism="Chlamydiaophila psittaci"	
Db	1021	CCGTGGAGTACTTATCTCAGTATCGAATCTCTGGAGACTTGGTCTTTCATGATCTGTG	1080	/strain="6BC"		/db_xref="taxon:83554"	
QY	1219	atccaagatacactccotttgggtgttaacagtaactcgaagctccctggtgagagatctgc	1278	/note="Specific host is the house parakeet, most probable species Melopsittacus undulatus."		568..831	
Db	1081	ATCCAAGATACACTCCCTTCTGTGTGTACAGTACTCGAAGCTCCTGGTGGAGAGATCTGC	1140	/gene="envA"		568..831	
QY	1279	tgtaataaagtgtgttgctgtataaagaatgtgccaggagaaacccctccagtttaa	1338	/gene="envA"		568..831	
Db	1141	TGTAATAAAGTGTGTCGCTATTAAAGAAATGTGCCAGGAGAAACCCCTCCAGTTAAA	1200	/codon_start=1		/transl_table=11	
QY	1339	ctgtagtgaaagctcaagttcctcggagattcacaaatacaattcagtaactagtgag	1398	/evidence=experimental		/product="low molecular weight cysteine-rich envelope protein"	
Db	1201	CTTGTAGTGAAGCTCAAGTTCTCTGGAAGATTACAAATCAAGTTGCAGTAAGTAGTGAG	1260	/protein_id="AAB61618.1"		/db_xref="GI:144490"	
QY	1399	tctaactcggacaatgtacatcttcgcgagaacaacaacacattggaaggtcttgca	1458	/translation="MKAVVLTATVCGVVLTSCCRIVDCCFEDPCAPKPCNPCGNK"		DKGCSGCVYTFSCSKPCGSECNPGVQGPQAKGCTSLDGRCKQ"	
Db	1261	TCTAACTCGGAACATGTACATCTTGCCAGAAACACACACATTTGAAAGGTCTTGCA	1320	997..2670		/gene="envB"	
QY	1459	gttaaccatattgctgtattagacacaaatgactctatctgtgtagagaaaatactgct	1518	997..2670		/gene="envB"	
Db	1321	GCTACCCATATGTGCTATTAGACAAAATGATCTATCTGTGTAGGAGAAATACATGCTC	1380	/codon_start=1		/transl_table=11	
QY	1519	tatcgtatctgtgtaactaacccgtgttctgtaagatacctaagctatctttaattctg	1578	/evidence=experimental		/product="high molecular weight cysteine-rich envelope protein"	
Db						/protein_id="AAB61619.1"	
QY						/db_xref="GI:144491"	
Db						/translation="MSKLIIRVVVLTALTSMASSFASGKTEAAAESLATRFRFTASTEN	
QY						SDDNVFOATAKKVRFGRNKNQRFQKHTGAFCDKBFYPCGGCQCPVDATQESCYCKM	
Db						YCVNRVNDNCNVETISQSVPEYATVGSPPYIELLAVGKKDCVNVVITQQLPCEVEFVSD	

```

PATTTSDSKLIWIDRLGQGEKCKIIVWPKPLKBCGCFPTAATVACAPBELSYTKCGQ
PAICIKGSPBCALRCPVKYKIEVCNTGSAIARNVVDNPDPDGTTHASGGQRLVSN
LGDMPGDSKFCVEFCQKRGKVNATVSYCGHKCSANVTYVNEPCVQVNIISGA
DWSYVCQEVYIIVSNPGLDLVDVIEDTAPSGATILEAAGAEICCNKAWCIKEM
CPGETLQFVKVAKAOSPGKFTNQVVKVNTSDSGCTSCAEVTHWKGLAATHMVCYIDT
NDPICVGNVYRICVTNRGSAEDTNVSLIKFSLKELQPVSSSGPKRGITGTGTVVFD
ALPKLGSKESEVFEFVTLKGIAPGDARGEAILLSSDLTPVPVADENTHY"
2877. 3281
/gene="srp"
CDS 2877..3281
/gene="srp"
/codon_start=1
/evidence=experimental
/product="sulfur-rich protein"
/protein_id="BAB61620.1"
/db_xref="GI:144492"
/translacion="MSGENANSIGSDVTSLIQPLEOVMODEGVQVSLINSVLGWCVRV
HLINPIKTSKIVOSRAFOITMVVLGIILLIAGLALTFVLQGLGKNFLFLIPAVIGL
VKLLATSCMEKPCPTPEKWRICKBLWQLKF"
BASE COUNT 1133 a 563 c 723 g 1131 t
ORIGIN

Query Match 56.1%; Score 1185.2; DB 1; Length 3650;
Best Local Similarity 77.3%; Pred. No. 7.7e-257;
Matches 1479; Conservative 0; Mismatches 428; Indels 7; Gaps 3;

Qy 1 ttgatcaggtagtagtaggaatgaattatctcgtacaccttaattca--gataataaac 58
Db 859 TTGTTAGGCAGTTAGGAGGTAATTTATTTCCCTGCTAACTGCTTTATGAAAAATAAAC 918

Qy 59 ccaaaatgtaggtaggaagttaacaaacattctaccgtagcagaagaaaaataa 118
Db 919 TTAATGTTGAGGGTAGAGTTAC--AACTTTTACCAGATGCGGAGGAGAAAAATAAC 976

Qy 119 acatgogataggagatccctatgtccaaactcatcagacgagtagtgatcgcttcgc 178
Db 977 ACAGCGATAGGAGATGCCATATGTCCAACTCATCAGACGAGTAGTTACGGTCTCCGCG 1036

Qy 179 taacgagtaggcgagttgcttggcagcggggtatagagcgccgtgtagcagagctc 238
Db 1037 TAACTAGTAGGGGAGTTTATTTGCCAGCGGGAAGATAGAGCGCCGCTGCTGCAGAGTCT 1096

Qy 239 tgattactaagatcgctgtagtcgga---caagccagcacctgttctctatgacag 295
Db 1097 TTGCTACAGATTCTATGCCAGTACTGAAACTCAGATGACAACTGTTTTCAGACACAG 1156

Qy 296 cgaagaaggtagactgtccgttagaataaacaaccaggtgaacaaaaagccgtagtg 355
Db 1157 CCAGAAAGTTAGATTGGTCGTAAACAAAATCAAAGACAGACAAACAAACATACTGGAG 1216

Qy 356 ctttttggataaagaattttatccctgtgaagaggagcagatgtcaacctgtagaggctc 415
Db 1217 CTTTCTGTGATAAAGAAATTTATPCTGTGGAAGGTGGCGAGTGCCCAACACGATAGACGCTA 1276

Qy 416 agcaagagttctgctacggaatgtattctgttaaaagtaaacgatatgtgcaactag 475
Db 1277 CACAAGAACTCTGCTACGCAAAATGATTGTGTCGGTTTAACGATGACTGTAAACGTG 1336

Qy 476 aaatttgcgcgtccgtccagaatacgtactgtgagatctccttaccctattgaaatcc 535
Db 1337 AAATACCCCAATCTGTACCTGAATATGCAACAGTAGAGTCTCCTTATCTATTGAAATTC 1396

Qy 536 ttgctataggcaaaaaagattgtgttgattgtgtattacacacagctacaccttgcgaag 595
Db 1397 TCGCTGTAGTAAAAAAGATTGCGTTAATGTTGTGATTACTCAACAGCTTCTCTGCGAAG 1456

Qy 596 ctgaatctgtaacgagtagcagaacaaactcctacaagtgaatggaaattagctcga 655
Db 1457 TTGAGTTGTACAGCAGTGTCTCGCACACACACCAACCTCGGATAGCAAAATTAATCGGA 1516

Qy 656 aaatcagatcgctgggtgcagagagataaattgcaaaattactgtatgggttaaaacctctta 715
```

```

Db 1517 CAATTGATCGCTTAGGTCAAGGTGAAAATGCAAAATACCGTTTGGGTAAAACCTCTTA 1576
Qy 716 aagaaggttgctgtctcaagctgctactgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 775
Db 1577 AAGAAGGTTGTGCTTACCGCGGCTACTGTATGTGTGTGTGTGTGTGTGTGTGTGTGT 1636
Qy 776 ctaaaatgcggccaacacgaccttctgtatgaagaagagcctgactgctgctgctaa 835
Db 1637 CCAATGCGGACACACGACTATTGTATTAAGCAAGAAGCTTCTGAATGCGTTCCTTAC 1696
Qy 836 gatgcctgtatgtctacaaaatcgaagtagtgaacacagagatctgtctatgtccgttaa 895
Db 1697 GTTGCCCAAGTTTGTACAAAATCGAAGTTTGAACACACAGGTTCTGTATTAAGCCGTA 1756
Qy 896 taactgtagataatctctgtcccgatggcttattctcgtcatctgtgtgtgtgtgtgtgt 955
Db 1757 TTGTCGTCGATAAACCCAGTTCCTCGATGGCTATCTCATGCTTCAGGACAAACGCGTCT 1816
Qy 956 ctttttaacttaggagacatgagacctgtgcgataaaaaggattattacagttgctgtgc 1015
Db 1817 CTTTTAACTTAGGAGATATGCTGCTGGGATTTCTAATGCTTCTGTGTGGAGTTTGGC 1876
Qy 1016 ctcaagaagaggtcaaatcactcaactgctgtctgtctgtctgtctgtgtgtgtgtgtgt 1075
Db 1877 CACAAAAGAGGAAAGATTTACCGTGGCTACTGTATCTTACTGCGGAGGACATAAAT 1936
Qy 1076 gtctgcgaaaatgtaactacagctgttaataaggcctgtgtacaaagtaaatatctctgt 1135
Db 1937 GTTCTGCTAACGTAACCTACTGTAGTTAACGAAACCATCGGTACAAATATCTCTGGAG 1996
Qy 1136 ctgattgtcttaccgtatgtataacctgtgagtagtctatctcagtagtcaatgctgtgtgt 1195
Db 1997 CTGACTGTGCTTATGTATGTATGATGCTGTAAAGCTGTAGAAATACACTATCTGTATC 2056
Qy 1196 acctgggttcttctatgtctgtatccaaagatacactcctcctgtgtgtgtgtgtgtgtgt 1255
Db 2057 ATCTTAAACTTTTACCATGATGATGTATAGAAATACCGCACCTTTCAGGAGCAACAAAT 2116
Qy 1256 aagctcctcgtgtgagagatctgctgtataaaggctgtgtgtgtgtgtgtgtgtgtgtgtgt 1315
Db 2117 AAGCTGCTGGAGCTGAAATCTGCTGTAAACAAAGCTGTATGTGTGCATCAAAAGAAAT 2176
Qy 1316 cagggaacacctccagctttaaactgtatgaagctcaagttccctgtgtgtgtgtgtgtgtgt 1375
Db 2177 CAGGAGAGACTCTTCAATTTAAAGTTGTGGCTAAAGCACAAAGCCAGGTAAATTCACAA 2236
Qy 1376 atcaagttgcagtaactagtagtcttaactgcggaacatgtacatctgtgcgagaacaa 1435
Db 2237 ATCAAGTTGTTGTCAAAACTAACCTCCGATTTGTGGAACTGTACTTCTTGTGCGAAGAT 2296
Qy 1436 caacacattgggaaggctctgcagctaccctaatagtcgctattagacacacaaatgacctta 1495
Db 2297 CAACACATTTGGAAGAGCTTTCGAGCTAGCATATGTGCTAATCGATACCAATGATGCTTA 2356
Qy 1496 tctgtgtaggagaaaatctgtctatcgtatctgtlaactaacgltgtgtctgtgtgtgtgt 1555
Db 2357 TTTGGGTAGGAGAAATACTGTATACCGTATTTGTGTAAACCAACCGTGTCTTGTGCAAG 2416
Qy 1556 atactaacgatacttcttaactctgtgaagttctcaaaagaacttcagacaaatagcttct 1615
Db 2417 ATACGAACGCTGCTGTAAATCTTAAAGTTTCTTAAGGAACATACACCCGTTTCTTCTTCAG 2476
Qy 1616 gtccaaactaaaggagacgtcttcaggtataacctgtgttttcgacgctttaccctaaactcg 1675
Db 2477 GTCCAACAAAAGAAACCATTTACAGGAATACAGTAGTATTTCTGATGCTTCTACCTTAAAT 2536
Qy 1676 gtcttaagaaactgtacagttcttctgttactgtgaaggtattgtctcccgagagagctc 1735
Db 2537 GTTCTAAGGAATCTGTAGATTTTCTGTAACTTAAAGGAATTTGCTCCAGAGAGATGCTC 2596
Qy 1736 gctggcgaagctatcttctctctatgatacactgacttccacagtagtaccagacagaaata 1795
```

[illegible]


```
DB 1107 TTGAGTTGTGACGAGTGATCTCTGCGACCAACACCAACCTCGGATAGCAAAATTAATCTGGA 1166
QY 656 aaatcgtatcgctgggtgcaggagataaaatgcaaaattactgtatgggtcaaaacctctta 715
DB 1167 CAATGTATGCTGCTTAGGTCAAGGTGCAAAATGCAAAATACCGTTTGGGTGAAACCTCTTA 1226
QY 716 aagaggttctgtcttcacagctactgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 775
DB 1227 AAGAAAGTTGTTGCTTACCGCGGCTACTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1286
QY 776 ctataatcggttcaaacaccatttatttaagaagaagaacctgaactgtctgtctgtctaa 835
DB 1287 CCAATGGGACCAACCAAGCTATTGTTATTAAAGCAAGAGTCTCTGAATGCGCTTGTCTAC 1346
QY 836 gatccctgtatgtctacaaatcgaagtgtgtgaacacaggtatctgtattgcccgtaacg 895
DB 1347 GTTGCCAGTTGTTGTACAAATCGAAGTTTGAACACAGAGTTCTGTATAGCCCGCTAATG 1406
QY 895 taactgtatgaataacctgtcccgatggtatctctcatgtatctgttcaaaagattctct 955
DB 1407 TTGTCGTGATACCCAGTTCCCGATGGCTACTATCTATCTATCTATCTATCTATCTATCT 1466
QY 956 ctittaatcttaggagacatgagaccctggcgataaaaaaggtattacagattgagttctgcc 1015
DB 1467 CTTTAACTTAGGAGATATCGCTCGGGATTTCTAAATGCTCTCTGTGTGTGTGTGTGTGTGT 1526
QY 1016 ctcaaaagaaggtcaaatcactaactggttctactgtactgtactgtactgtactgtactgt 1075
DB 1527 CACAAAAAGAGGAAAGTTTACTAAGCTGGCTACTGTATCTTACTTGGCGAGGACATAAT 1586
QY 1076 gtctgcaaatgtcaactcacagttgttaagtgcctgtgtgtgtgtgtgtgtgtgtgtgtgt 1135
DB 1587 GTTCTGCTAAGTAACACTACTGTATTAACGAACCACTGCTCAACAGTAATAATCTCTGGAG 1646
QY 1136 ctgattgtcttaagtgttaaacctgtgagtaactctctcagatcagataactgagag 1195
DB 1647 CTGACGTGCTTATGTATGTAAGCTGTAGTAATCACTATCTGTATCTTAACCCAGAG 1706
QY 1196 actgtgttcttcattgtatgctgatacgaagatacacctcctctgtgtgtgtgtgtgtgtgt 1255
DB 1707 ATCTTAAACTTTAGCATGTAGTTATAGAAGATACCCGACCTTCAGGAGCAACAATTTTAG 1766
QY 1256 aagctccgtgtgagagatctgtgtaataaagttgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1315
DB 1767 AAGCTGCTGGAGCTGAATCTGCTGTAAACAAAGCTGTATGTGTGATCAAGAAATGTGCC 1836
QY 1316 caggagaaacctccacttcaactctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1375
DB 1827 CAGGAGAGCTCTTCAATTTAAAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1886
QY 1376 atcaagtgcagtaactagttagtcttaactggtggaacatgtacatttgcgcaaaacaa 1435
DB 1887 ATCAAGTTGTTGTCAAAACTAACTCCGATTTGTGGAACCTGTACTTCTTCCGCAAGATTA 1946
QY 1436 caacacattgaaaggtcttgacagtlaccatattgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1495
DB 1947 CAACACATTGGAAGGCTTTCAGCTAGCATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2006
QY 1496 tctgtgtgagaaataactgtctatctgtatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1555
DB 2007 TTTGCTGAGGAAATACTGTATACCGTATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2066
QY 1556 atactaactatctttaaacttgaagtcttcaaaagaacttcagacaaatagtcttcttaag 1615
DB 2067 ATACAAAGCTCTCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 2126
QY 1616 gtccaaactaaaggaacttccagtaataaccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1675
DB 2127 GTCCAAACAAAGGAACCAATTACAGGGAATACAGTAGTATTCGATGCTCTTACTTAAATTA 2186
QY 1676 gtctaaagaactgtagaatttctgttacctgaaggttattgtctcccgagatgctc 1735
DB 2187 GTTCTAAGGAACTGTAGAGTTTCTGTAACTTAAAGAAATTTGCTCCAGGAGATGCTC 2246
```

```
QY 1736 gcgcgagagctattcttcttctgtatcacactgactccagctatcacagacacacaaata 1795
DB 2247 GAGGAGAGGCTATTCTTCTTTCAGACACTTAAACGTTACCTTCTTCTGCTGATACAGAAACA 2306
QY 1796 ccacgctgtattaaattcctaagaataattcctaagaagcagcagatattccgctctgttt 1855
DB 2307 CACATGTTTATTAATCTTTTAAATTTTCTTAAAGATTGTTAAACGAAAGGCGGT 2366
QY 1856 aggtatgcttcaaaagtcacgctttagctacacttactactactaaagcggtttt 1909
DB 2367 CTTAAGATCTTAGGAGGCTTTTGTGTTCTTCTTCTATCTGCTACGATTTTCT 2420

RESULT 9
CPU41759 8950 bp DNA linear BCT 07-MAY-1999
LOCUS Chlamydia caviae strain Guinea Pig Inclusion Conjunctivitis
DEFINITION (GPIC) RecJ recombination protein gene, partial cds; glutamyl-tRNA
synthetase homolog, outer membrane protein 3 (omp3), outer membrane
protein 2 (omp2), and hypothetical sulfur-rich protein (srp) genes,
complete cds; and unknown genes.
ACCESSION U41759.1 GI:1783376
VERSION U41759.1
KEYWORDS cysteine rich protein; outer membrane protein; glutamyl-tRNA
synthetase.
SOURCE Chlamydia caviae.
ORGANISM Chlamydia caviae.
REFERENCE 1 (bases 1 to 6000)
AUTHORS Hsia, R.C. and Bavoull, P.M.
TITLE Homologs of Escherichia coli recJ, gtx and of a putative 'early'
gene of avian Chlamydia psittaci are located upstream of the 'late'
omp2 locus of Chlamydia psittaci strain guinea pig inclusion
conjunctivitis
JOURNAL Gene 176 (1-2), 163-169 (1996)
MEDLINE 97075925
REFERENCE 2 (bases 6001 to 8950)
AUTHORS Hsia, R.C. and Bavoull, P.M.
TITLE Sequence analysis of the omp2 region of Chlamydia psittaci strain
GPIC: structural and functional implications
JOURNAL Gene 176 (1-2), 155-162 (1996)
MEDLINE 97075924
REFERENCE 3 (bases 1 to 8950)
AUTHORS Hsia, R.-C. and Bavoull, P.M.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-1995) Ru-ching Hsia, Microbiology & Immunology,
University of Rochester, 601 Elmwood Avenue, Box 672, Rochester, NY
14642, USA
FEATURES
source
1..8950
/organism="Chlamydia caviae"
/strain="Guinea Pig Inclusion Conjunctivitis (GPIC)"
/species="guinea pig"
/db_xref="taxon:83557"
/notes="submitted as Chlamydia psittaci"
<1..809
/notes="ORF1: homolog of Escherichia coli recJ"
/codon_start=3
/transl_table=11
/product="RecJ recombination protein"
/protein_id="AAB4138.1"
/db_xref="GI:1783377"

CDS
1250..2404
/notes="ORF2"
/codon_start=1
/transl_table=11
/product="unknown"

CDS
```

```
/protein_id="AAB41139.1"
/db_xref="GI:1783378"
/translation="NFSVQSSVIGNSLITMNPQETLIFSSFAARQAHRRRTLPVI
YKILVDICIKICIPRIILIFPGFLVGLKICONCLLPAAAGLFTFOLCFPAKKILOE
SPFYVDWLESSEYASSVKRIPIQCDOLLIDAISOPPAEDKRMALVSLGNSCFET
RAWYSRNPWILNVAQKARAVVDFYPCYHSHKGPITRESLCKAQCYVHLRHPD
KAMQDILATIGLTYQALGKREYVDSGDGVDMFVFKDRGPRASVCAIASGLWGRV
AGWTATGWEINSAISYSELVCFTHGEDTQAQIGDMFNDRNCFAAPFLDPNA
PFPFGKKIPVGEFLQHKGEUDETIIQNVVKKHIDHYSPEPSPETPDGE"
3005..3550
/notes="ORF4; homolog of EUO of Chlamydia psittaci 69C"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAB41140.1"
/db_xref="GI:1783379"
/translation="NECIQHESCFDLDRDNAQOLEVOEGTEMVSTOAAKLHNVTRO
AIYVAJKOKKLKASKTTRWEIDLKDEYKRNYSRKSLYOGELFDNDKGFYSVNO
VADMLGIPQKVYATRTGTRMGERKGAANVISQSELDIRYKSEYLNKQIVKKKEVAV
EHPATPADAVESGTALFEND"
3822..5339
/notes="ORF5; homolog of glx of prokaryotes"
/codon_start=1
/transl_table=11
/product="glutamy1-trna synthetase homolog"
/protein_id="AAB41141.1"
/db_xref="GI:1783380"
/translation="NAWENVRVRVAPSGTPDPHVGTAYMALFNEIFAKRFNGKMILRI
EDTQTRDDYDEKNIIFSALQGIQWDEGPDIGGPHYRQSRTEIYREYAEILLK
TDYAYKCFATPELEMAVATTLGIRGGYDRRYLSPSEIEARTQEGQPTIRLKV
PLTGEEDYCKYGVFPVPMADVDQVLMKSDGEPYHFRANVDDHLMGITHLVREE
WLSSTPHLLLYFAFGWPPPIFLHMPDLLNDGDTKLSKRNPTSIIFYRDACYKEAF
NPLTLWGLSMGDEFEVYSLKLANEDPKICKSGAVSDVRKLDMNKHLYNHGSP
ENLLALBKLDYNDRELLKILPLCSRMALASPVGLSEPFESVLPYKSEELPAAI
SQKAILYSYVYKLEKTDLVNDQFYLAKSLSEAFQVHHKKVYIPLLYVAITGK
QGFLPDSMELUGKPRTRRWVHQNLGLGVPKAIQIADIKLVKEENFENKILFE"
5992..6258
/gene="omp3"
5992..6258
/gene="omp3"
/codon_start=1
/transl_table=11
/product="outer membrane protein 3"
/protein_id="AAB41142.1"
/db_xref="GI:1783381"
/translation="MKAILLAACVGVGLSSCCRTVDCCFEDPCAPKSCNPVEX
KDGECSGCSYTPSCSKPGGSDCAKVGQPAKGCITSPDGRCKQ"
6423..8099
/gene="omp2"
6423..8099
/gene="omp2"
/notes="developmentally-regulated; cysteine-rich"
/codon_start=1
/transl_table=11
/product="outer membrane protein 2"
/protein_id="AAB41143.1"
/db_xref="GI:1783382"
/translation="MSKLFRRVTVLTALTSMASSFASGKTEAAABESLATREIASAEN
SDNISORAKKVRGRNKKYOKQKNGNPYCKDEFPYCODGSCQSSVDTKQPCYVK
MVCYVSDSNVEISOAPEVATVGSVPRIELAVGKDCVNVVITQOLPCREVEILSS
DPATPTPSKLIWPIIDBLGGERCKITVMWPLKCEGCTAATVACPELSYKCG
QPAICIKQSPCACLRCPVYKYLEVNTGSAIARGVVDNVPDQISRPSQORVLSF
NLDMPGDKCTVEFCQKRGKITNVATVYSCGHKCSANVTIVINPEQVNVISG
NMCEYLOQKPEVETIYVSNPDKLIDYVIEDLIPSGITLLEAPGAEICCNKAWCIE
MCPEYLOQKPEVETIYVSNPDKLIDYVIEDLIPSGITLLEAPGAEICCNKAWCIE
TNDPICVGNENTYRIVCTNRGSAEDTNVSLILFKSELOPVSQSGPKTKITGTWVWF
DALPKLGSSESYEFVTLKGVAPDARGEAILSSDITLVFPVADTENTHY"
8310..8792
/gene="srp"
8310..8792
/gene="srp"
/codon_start=1
/transl_table=11
```

CDS

CDS

gene

CDS

gene

CDS

gene

CDS

```
/product="hypothetical sulfur-rich protein"
/protein_id="AAB41144.1"
/db_xref="GI:1783383"
/translation="MLTGVNSESSEGVVIDLIKPLDQVAKMNETVOVTLVNSVLGCKAH
IVDPIKSKIVQSFQFOITMVVLGVILLTAGLALFVLOGLGKNALFLIPAVIGLV
KLLTTSVPNEKPCFPEKWRCKRLALTATEDLDDQIQNSNTLFTTSSSDVTNTATOS
"
BASE COUNT 2724 a 1612 c 1858 g 2756 t
ORIGIN

Query Match 54.9%; Score 1158.8; DB 1; Length 8950;
Best Local Similarity 78.4%; Pred. No. 6.2e-251;
Matches 1440; Conservative 0; Mismatches 387; Indels 9; Gaps 4;

Qy 1 ttgtcaggtagttaggagatgaataattctgaacttaattca-gataataaac 59
Db 6286 TTGGTTAGGTAGTAGGAGGTAAATATTTCCTGCTTAACTACCTTTATCAAAATAACT 6345
Qy 60 caaatgttgagggttaagagtttcaaaaacattctaccgatggcagagaaaaataaa 119
Db 6346 TAAATGTTAGGTTAGAGTCAAGTCC--CATCATTTCACCCGATGCGAAGAAAATAACR 6403
Qy 120 catgcgataggagatccctatgtccaaactcatcagacgagtagttacggtccttcgct 179
Db 6404 CACCGATAGGAGATCCCTATGTCCTCAAACTCTTCAGACGAGTAGTTACGGTCTCCGCT 6463
Qy 180 aacgagttatggcaggttgcttttgcagcgggggtatagagccgagctatagcagtgctct 239
Db 6464 AACTAGTATGGCGAGTTTCATTTCGCGCGCAAGATAGAGGCCCTGCTGCACAGTCTCT 6523
Qy 240 gattactaagatcgctcgtatgtcgagaaa---caaacgcagccctgttctcctagcagc 296
Db 6524 TGTCTACAAGATTATTGCTAGTGTCTGAAAACCTCAGATGACNATATTTCGCAGAGAAAGC 6583
Qy 297 gaagaggttagactgttcgtagaataaaacaacccaggtgaacacaaaacgcgtagtgctc 356
Db 6584 CAAGAAGTTAGATTTCGTCGTAATAATAATCAAAAACAAGAACAAAAAATAATCGCC 6643
Qy 357 ttttgtataaagaattttatccctgtgagaggggagcgatgtca---acctgtagagcg 413
Db 6644 TTACTGTGATAAAGAAATTTTATTCCTGCTCAAGATCGGCTCATGCCAATCATCATAGTAC 6703
Qy 414 tcagcaagagctctgtcagcgaagattgtattctgtaaaaagataaacgagtagttgcaagct 473
Db 6704 TAAGCAAGAACCCTTGCTCGGCAAAATGCTATTGCTGCTAGCGATGATAGTAACCT 6763
Qy 474 agaaatttccagtcogttccagaatacgcactactgtaggatctccttaccctattgaaat 533
Db 6764 TGAATTTAGCCACAGCTGTACCTTGAATATGCAACAGTAGGATCTCTTATTCCTATTGAAT 6823
Qy 534 ccttgctataggcaaaaaagattgctgtgagtgtgattatcacacacagctacactgcaga 593
Db 6824 TCTTGCTGTAGGTAAAGAGATTGCGTTTAAATGTTTATCATCAACAGCTTCCTTTCGCA 6883
Qy 594 agctgaattcgttaagcagtgatccagagaaacccctcacagtgatgggaaattagctctg 653
Db 6884 AGTTGAATTCATAAGCAGCGATCCTGCGACAAACCTACTTCCCGATAGTAATAATTTG 6943
Qy 654 gaaattcgcactcgttgtagagagataaataatcactgtatgggtgaaacccct 713
Db 6944 GACAATTGATCGCTTASGTCAAGGTGAAGATGCAAAATTTACTGTTTGGGTAAACCTCT 7003
Qy 714 taagaaggttgctgctcagcagctgctactgtatgtgcttgccagagctccgcttcta 773
Db 7004 TAAAGAAGGTGCTGCTGCACAGCTGCTACCGTATGCTTGGCCAGAACTTCGCTCTTA 7063
Qy 774 tactaatgcgggtcaaccagccatttgtattaaagcaagaaggacactgactgtcttgct 833
Db 7064 TACCAATCGGCAACACCTGCTATTTCGATTTAAGCAAGAAGGCCCTGAGTGCCTTTCGCT 7123
Qy 834 aagatgcctctgtatgctacaaatcgaagtagtgaacacaggaactgctattcccgtaa 893
```


Query Match 54.8%; Score 1156.6; DB 1; Length 3186; Best Local Similarity 78.3%; Pred. No. 2.2e-250; Matches 1427; Conservative 0; Mismatches 389; Indels 7; Gaps 3;									
QY	1	ttgatcaggttagtgagagatgaatataattctctgactacttaattca--gataataaac	58						
DB	391	TTGGTTAGGTAGTATAGGAGGTAAAGTATTTCCCTCTAATCGCCTTTATGAAAAATAAC	450						
QY	59	ccaaatgttgagggtaagagtttacaacaacattctaccggatgagcagaagaaaaataa	118						
DB	451	TTAAATGTTGAGGGTAAGAGTTTAC--AATTTCTACCCGATGCGAGAAGAAAAATAAC	508						
QY	119	acatcgatagagatccctatgtccaaactcatcagacgagtaggttacggtccttcgc	178						
DB	509	ACACGGATAGGAGATCCCTATGTCCAAACTCATCAGACGAGTAGTACGGCTCTCGCG	568						
QY	179	taacgagttgagcagttgtcttgcagcgggggtatagagccgctgtgacgagctc	238						
DB	569	TAACTAGTATGGCGAGTTCAATTTGCCAGCGGAGATAGAGCGCGCTGCTGCAGAGTCTC	628						
QY	239	tgattactaagatcgctcgtagtcggaaa--caaacgcagccctgttctctatgacg	295						
DB	629	TTGCTACAAAGATTCATTTGCCAGTACCGAAAACTCAATGACAATGTTTTCGAACAACAG	688						
QY	296	cgaaaggttagaactgtccgtatagaataataacaaccagttgaacaaaaagccgtgtg	355						
DB	689	CCAAGAAAGTTAGATTGTCTCTACAAAAATCAAAGACAAGACAACAAACATACTCGCG	748						
QY	356	cttttctgtataaagaattttatccctgtgaagaggggcagatgtcaacctgtagagctc	415						
DB	749	CTTTCTCTGATAAAGAAATTTTATCTTCGGAAGGTGGTCAGTGGCAATCCGTCGATCTA	808						
QY	416	agcaagctcttgcacgaagatgtattctgttaaaagtaaacgagatgctgcacatag	475						
DB	809	CACAAGAACTTGTACGSCANAATGTATGTGTCCGTGTTAACGATGACGTAAACGTGG	868						
QY	476	aaatttgcagctcgttcagaaatcagctactgttaggtctctctaccctattgaatcc	535						
DB	869	AAATTAGCAAGCTGTACCTGAATATGCAACAGTAGGATCTCTATCTCTATTTGAAATTC	928						
QY	536	ttgctatagcgaataaagattgtgtgtgtgtgtgattacacacacagctacacctgcgaag	595						
DB	929	TTGCTGTAGTAAAAAAGATTCGGTAAATGTGTGTGATCCTCAACAACCTTCTTGGCAAG	988						
QY	596	ctgaatctgaagcagtgatcagaaacaactcctacagtgatgggaataatagctcga	655						
DB	989	TTGAGTTTGCAGCAGTGATCTCGCAACAACACCAACCTCAGATAGCAAAATTAATCTGA	1048						
QY	656	aatcgtatcgcctgggtcaggagataaaatgcaaaattactgtatgggtaaaaacctetta	715						
DB	1049	CAATTGATTCCTTAGTGTCAAGGTGAATAATGCAAAATTAACCGTTTGGGTAAACCTCTTA	1108						
QY	716	aagaagtttgcgtcttcacagctgtactgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgt	775						
DB	1109	AAGAAGGTTTGTCTTACCGCGGGCTACTGTATGGCGCTTGGCCAGAACTTGCCTCTATA	1168						
QY	776	ctaaatcggttcaacagccatttattatagcaagaaggaactgactgtcttgcctaa	835						
DB	1169	CCAAATCGCGACACACAGCTATTTGTATTAAGCAAGAAGCGCCCTGAGTGTCTTGCCTAC	1228						
QY	836	gatgcctgtatgctacaaaaatcgaagtagtgaacacaggaatctgtctattgtccgtaacg	895						
DB	1229	GTTCGCCAGTTTGTACAAAATCGAAGTTTGCACACACAGGTCTGTCTATACCCCGTAATG	1288						
QY	896	taactgtatataatccctgtcccgatggtctattctcatgcatctgtgtcaaaagattctct	955						
DB	1289	TTGCTGTGATACCCCTGTTCCTCGATGGCTATATCTACCTACCGCTTCAGGACAACCGCTTCTT	1348						
QY	956	cttttaacttagagacatgacactgcgcataaaaaggtatttacagttgagttctgcc	1015						
DB	1349	CTTTAACTTAGGAGATATGCCCTCGGGATTTCTAAATGCTTCTCTGTGAGTTTGGCC	1408						
QY	1016	ctcaagaagaaggtcaaatcaactaacctgttgcactgtactgtactgtactgtactgtactgt	1075						

DB	1409	CGCAAAAAGAGGAAAAAATTACTAACCTAGCTACCGTATCTTCTCGGAGGACATAAAT	1468
QY	1076	gttctgcaaatgtaactacagcttgaatgaagcttgtgtacaagtaaatatctctgggtg	1135
DB	1469	GTTCGCCGAAGCTAACTACTGTAGTTAACGAACCTCGGTACCAAGTCNATATCTCTGGAG	1528
QY	1136	ctgatttgcttaogtatgaacctgtgaggtactctatctcagtatcgaatccctggag	1195
DB	1529	CTGACTGGTCTTATGTATGTAAAGCTGTAGAATACACTATCGTTGTGTCCAACCTAGGGG	1588
QY	1196	acttggttctcatgactgctgacccaagatatacactccctctctggtgttacagtaactcg	1255
DB	1589	ATCTTAAACTTTACGATGCTCTGTAGAGATACCGTACCTTCAGGAGCTACAAATTTTAG	1648
QY	1256	aagctccctgggtggagagactgctgtataaaagtgttctggcgattataaagaagaatgtgc	1315
DB	1649	AAGCCGAGGAGAGCTGAATCTGCTGTAAACAAAGCTGTATGGTGCATCAAGAGATGTGCC	1708
QY	1316	cagggaacacccctccagtttaaaactgttagtgaagctcaagtctcctgggaagattcacaa	1375
DB	1709	CAGGAGACACTCTACAATTTAAAGTTGTCTAAAGCACAAAGCCCAAGTAAATTCACAA	1768
QY	1376	atcaagtttgcagtaactagttgacttaactgcgaacatgtacatcttgcgcaaaacaa	1435
DB	1769	ATCAAGTTGTGTCTNAAACAACTCCGATTTGTGGNACTTGCATCTTGTGTGCACAAAGCTA	1828
QY	1436	caacacattggaaaggtctctgcagctaccatatactgctgattagacacacaaatgtacct	1495
DB	1829	CAACCAATGGAAAGGTCTGGCAGCTACTCATATGTGCTGAATGCATACCAATGATCTCTA	1888
QY	1496	ctctgttaggaagaaaaactctctatctgtatctgtgtaactaaacgtgtgtctctgtgaag	1555
DB	1889	TTTGGCTAGGAGAAATACTGTATACCGTATTTCTGTAAACAAACCGTGTCTTCACAGAG	1948
QY	1556	atactaaagatctttaaactgtgaagttctcaaaagaacttcagcaaatagcttcttcag	1615
DB	1949	ATACTAAGCTTTCTAATTAATCTGTAAGTTTCTAAGGAGCTGCAACCCGTTTCTTCTTCAG	2008
QY	1616	gtccaaactaaaggaagcattccaggttaataccgtgtgtttctgcagcgttttaccctaaactcg	1675
DB	2009	GCCCAACAAAAGAACCAATACAGGAANAATACAGTCTGTATTCGATGCTCTGCCCTAAATTAG	2068
QY	1676	gtcttaagaatctgttagagtttctctgttacctgtgaaggtattgtctcccgagagatgctc	1735
DB	2069	GTCTTAAGSAATCTCTACAGTTTCTCTGTAAACATTTAAAGGGATTGCCCTGGAGATGCAC	2128
QY	1736	ggcggaagctattcttctctctgatacactgacttcaccagtatcagacacagaaaaata	1795
DB	2129	GAGGAGAGCTATCTTCTTTCAGATACCTTTAACGGTACCCTGTTGCTGTACAGAAATA	2188
QY	1796	ccacgtgtattaaattcttaagg	1818
DB	2189	CGCAGGTTTATTAATCTTTTACG	2211

RESULT 11

LOCUS	CPU76760	1604 bp	DNA	linear	NC_13-MAY-1999
DEFINITION	Chlamydomophila abortus cysteine-rich outer membrane protein	Omp-2			
ACCESSION	U76760				
VERSION	U76760.1	GI:4098314			
KEYWORDS					
SOURCE	Chlamydomophila abortus.				
ORGANISM	Chlamydomophila abortus				
REFERENCE	Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.				
AUTHORS	Sheehy, N., Markey, B. and Quinn, P. J.				
TITLE	Sequence analysis of C. psittaci and C. pecorum 60 kDa genes				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1604)				
AUTHORS	Sheehy, N., Markey, B. and Quinn, P. J.				


```

QY 1601 caatagcttctcaggtccaactaaaggaacgatttcaggtaataacacggtgttttcgacg 1660
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1441 CTATATCTCTCTGGCCCAACCAAGGAACAATTTCTGGAATACAGTAATTTTTCATG 1500
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1661 ctttacctaactcgtcttaaggaatctgtagattttctgttacccttgaagaagttattg 1720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1501 CCTTACCAATATAGGTTCAAGAGATCTGTAGAGTTTCTGTACGCTTGAAGGGGATTG 1560
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1721 ctccggagatgtctcggcgaagctattcttctctctgataca 1764
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1561 CTCCTGGAGATGCTCGAGGAGAGGCGTGTCTTTTCTGTGATACA 1604
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
CHTOMPE
LOCUS CHTOMPE 2390 bp DNA linear BCT 26-APR-1993
DEFINITION Chlamydia trachomatis outer membrane protein(s) (omp2A and omp2B)
            genes, complete cds.
ACCESSION M85197
VERSION M85197.1 GI:144604
KEYWORDS outer membrane protein.
SOURCE Chlamydia trachomatis
ORGANISM Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE 1 (bases 1 to 2390)
AUTHORS Zhang, Y.-X. and Caldwell, H.D.
TITILE The nucleotide sequences of 10 and 60 kDa cysteine-rich outer
        membrane protein genes of Chlamydia trachomatis serovar F
JOURNAL Unpublished
FEATURES
    source          Location/Qualifiers
        1..2390
            /organism="Chlamydia trachomatis"
            /strain="TW-3/OT serovar C"
            /db_xref="taxon:813"
            /genes="omp2A"
            /genes="omp2B"
            /codon_start=1
            /transl_table=11
            /product="outer membrane protein"
            /protein_id="AA23158.1"
            /db_xref="GI:144605"
            /translation="MKKTAALLAALCSVLSLSCRIYDCCFEDPCAPIQCSPECSSKK
            DVGGCNSGIVPACPKCGGDTGDAKHGPAQARGIPVDKCRQ"
            /genes="omp2B"
            /genes="omp2B"
            /genes="omp2B"
            /codon_start=1
            /transl_table=11
            /product="outer membrane protein"
            /protein_id="AA23159.1"
            /db_xref="GI:144606"
            /translation="MNLKLRRAVITFAVTSVASFASGVLETSMAESLSNIVTSIADT
            KAKDNTSHSKARKNHSKETPDVRKEVPAPVHESKATGPKDSCFGRMVTVKVNDDRN
            VEITQAPSVATVGSPIETIATGKDCVDVLIITQOLPCEAEFVSDPATPTADGK
            LVNKIDRLQGGKSKITVWVKPLKGGCFTAAATVCAGPEIRSVTKGQQAICVKQEGP
            ENACRCPVYVYKINVNQGTATARNVVENPVDGYAHSSGQVRLVTLGDMQPGEHR
            TIVFECPLKGRATNIATVSYCGHKNTASVTVIINPCQVQVSIAGADWSYVCKPVE
            YVTSNPNGLDLRDVYVEDTSLPGVTVLEAAGAOISCNKVVTVKELNPGESLOKYV
            LVAQTPGQFTNNVVKSCSDGCTSCAEATYKGVAAATHMVCVVDTPVCVCGENT
            VBYICVTNGSLEDNTNLSMLKFSKELQPVFSFGPTKGTITGNTVVPDLSLPRLSKET
            VESVPLKAVSGDARGENALISSDPLTPVPSDENTHIY"
BASE COUNT 711 a 439 c 526 g 714 t
ORIGIN

Query Match 42.1%; Score 889.6; DB 1; Length 2390;
Best Local Similarity 71.0%; Pred. No. 3.4e-190;
Matches 1249; Conservative 0; Mismatches 479; Indels 32; Gaps 4;

```

```

QY 61 aaatgttaggggtaa----gagttacaaacatctaccgcatggcagaagaaaaaa 116
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 555 AAATGTTGAGGGTAAAGTTAGTTAATAACAATTTCTACCCGATGGCAGACAAATAATAA 614
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 117 aaacatgcgtagagatccctatgtccaaactccatcagacgagtagttacggtccttgc 176
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 615 TCTATGCGAATAGAGAT--CCTATGAACAACACTCATCAGACGACGAGTGACGATCTTCGC 673
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 177 gctaacgagtatggcgagttgttccagcggggggtatagaggcgcgtgtagcagagtc 236
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 674 GGTGACTAGTGTGGCGAGTTTATTTCTAGCGGGGGTGTAGAGACCTCTATGGCAGAGTC 733
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 237 tctgattactaagatcgctcgctagtcggaagaacaaagccagcacctgttctctatgacgc 296
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 734 TCTCTCTACAAACGCTTATTAGCTTAGCTGACACCAA-----AGC 772
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 297 gaagaaggtttagactgtccgtagaataaacaaccagttgaacaaaagacccgttggtgc 356
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 773 GAAAGCAACACACTCTCTATATAAAGCAAAAGCAAGAAAAACCAACACCAAGAGACTCC 832
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 357 tttttgtataaagaattttatccctgtgaagagggaagatgtcaacctgttagagggtcca 416
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 833 CGTAGACCGTAAAGAGGTTGTCTCCGGTTCATGAG-----TCTAAGGCTACAGGACCTAA 886
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 417 gcaagagcttctgctacggaagattgtattctgttaaaagtaaacgtagtattgcaacgtaga 476
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 887 ACAGGATTTCTTGCTTTGGCAGAAATGTATACAGTCAAAAGTTAATGATGATGATGATGATGA 946
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 477 aattggcagtcggttccagaatacgtactgttagatctcttaccctattgaaatcct 536
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 947 AATCACACAACCTGTCTCTGAATATGCTACGGTAGAGATCTCCCTATCTTATTTGAAATAC 1006
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 537 tgcattaggccaaaagaattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 596
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1007 TGCTACAGGTAAGAGGATTTGTTGTATGTTATCATCTACTAGCAATATACCATTGTGAAGC 1066
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 597 tgaattcgttaagcagtgatccagaacaaacctcctcaagtgatgggaaatlagttcgaa 656
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1067 AGAGTTCTGTACCGAGTGTATCCAGCGACAACCTCTACTGCTGATGATGATGATGATGATG 1126
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 657 aatcgatcgtcgtgtgcagagataataatgcaaaattactgtgtgtgtgtgtgtgtgtgtgtgtgt 716
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1127 AATGACCGCTTAGACACAGCGCAAGAGAGTAAATTAATGATGATGATGATGATGATGATGATG 1186
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 717 agaggtgtcgtcgtcagctactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 776
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1187 AGAAGGTGTGCTTTTACAGCTGCAACAGTATGCGCTTGTCCAGAGATCCGTTGCGTTTAC 1246
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 777 taaatgcggtcaacacgaccttgtattgaagaagagagacacctgactgtgcttgcctaaag 836
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1247 AAATGTGGACAACCTGCTATCTGTGTTAAACAAGAAGGCCAGAGAAATGCTTTGTTGCG 1306
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 837 atgcctgtgtgtgtacaaatcgaagtagtgaacacagactgtctattgcctgaacgt 896
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1307 TTGCCAGTAGTTTACAAATTAATGTAGTGAACCAAGGAACAGCAACAGCTGTAACGT 1366
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 897 aactgtagataactcctgttccogattggttattctcatgcatctgtgtcaaaagagttctctc 956
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1367 TGTTGTTGAAATCCTGTTCCAGATGTTTACCGCTTACTTCTTGACACGCGGTACTGAC 1426
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 957 tttlaacttaggacatgacgtacctcgcgataaaaggattattacagttgagttcttgcct 1016
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1427 GTTTACTCTTGGAGATATGCAACCTGAGAGACACAGCAAAATTAATGATGATGATGATGATG 1486
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1017 tcaagaagaggttcaaatcaactaactgttactgtactgtactgtactgtactgtactgtactgt 1076
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1487 GCTTAAACGTGGTGTCTACCAATATAGCAACGCTTTCTTACTGTGGAGGACATAAAA 1546
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1077 tctgcgaatgtaactacagttgttaatgagccttgtgtacaaagtaaatatctctgtgtgc 1136
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1547 TACAGCAAGCGTAACAACCTGTGTATCAACGAGCGCTTCGCTACAAGTAAGTATTATTTGCAGSAGC 1606
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```
QY 1137 tgattgcttcacgtatgtataaacctgtggagtagctctctctcagtagatcgaaatctggaga 1196
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1607 AGATTGCTCTTATGTTGTAAGACCTGTAGAATATGCTGATCTCCGTTTCCAATCTCGAGA 1666
QY 1197 ctgtggtcttcacgtatgtatcaagaatacacatccctctctgtgtgtacacgtactcga 1256
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1667 TCTTGTCTTCGAGATGTCGTCTTGAAGACACTCTTTTCTCCGAGTCAAGTCTCTGA 1726
QY 1257 agctccctgtggagagatctgtgtaataaagtcttggcggtatttaaaagaatgtgcc 1316
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1727 AGCTGCAGGAGCTCAATATCTTGTATTAAGACTAGTTTGGACTGTGAAGAACTGAATCC 1786
QY 1317 agggagaacccctccagtttaactgttagtgaagctcaagtctcctgtgaagattccaaaa 1376
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1787 TGGAGAGTCTCTACAGATAAAGTTCTTAGTAAGAGACAAACTCTCTGACAAATTCACAA 1846
QY 1377 tcaagttgcagtaactagtagtctaaactgcggaacatgtacatcttgcgcagaacaac 1436
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1847 TAATGTTGTGTGAAGAGCTGCTCTGACTGTGGTACTTGTACTTCTTGCAGAACGAC 1906
QY 1437 aacacattgaaaggctcttcagctaccatattgctgtattagacacaaatgatctcat 1496
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1907 AACTTACTGAAAGGAGTTGCTGCTACTCATATGTGCTAGTACTACTTGTGACCTGT 1966
QY 1497 ctgtgtaggagaaaaactgtctatctgtctgtgttaactaaacgtgtgtctgtgaaga 1556
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1967 TTGTGTAGGAGAAATACTGTTTACCCTATTTGTCTCACCACAGAGGTTCTGCAGAGA 2026
QY 1557 tactaactgtattttaaattgttaagttctcaaaagaacttcagccaatagcttcttcagg 1616
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2037 TACAATGTTTCTTAAATGCTTAAATCTCTAAAGAACTGCAAGCTGTATCTCTCTGG 2086
QY 1617 tccaactaaaggagacatttcaggtaataccgtgttttctgcagcttaccataaactcgg 1676
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2087 ACCAACTAAAGGAACGATTACAGGCAATACAGTAGTATTCGATTCTTACCTAGATTAG 2146
QY 1677 ttctaaagaaactgtagattttcttaccctgaaggtattctcccggaatctcgt 1736
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2147 TTCTAAGAAACTGTAGAGTTTCTGTAACTTGAACATGGAACGATACAGCTGGAATGCTG 2206
QY 1737 cggcgaagctattcttcttcttgatacacactgacttcacacagatcacagacacagaaatac 1796
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2207 TGGGGAAGCAGATTCTTCTTCGATACATATGACTGTTCCAGTTTCTGATACAGAGAATAC 2266
QY 1797 ccagctgtattataaattctaa 1816
||||| ||| ||| ||| |||
Db 2267 ACACATCTATTAATCTTTGA 2286
```

```
RESULT 14
CTCROMP CTCROMP 2385 bp DNA linear BCT 02-FEB-1993
LOCUS CHlamydia 60kDa cysteine rich outer membrane protein gene.
DEFINITION X53510.1
ACCESSION X53510.1 GI:40681
VERSION outer membrane protein.
KEYWORDS Chlamydia trachomatis.
SOURCE Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
ORGANISM 1 (bases 1 to 2385)
REFERENCE Watson, M.W.
AUTHORS Direct Submission
TITLE Submitted (22-JUN-1990) Watson M.W., Southampton University England,
JOURNAL Department of Microbiology, Faculty of Medicine, South Lab and Path
block, Southampton General Hospital, Tremona road, Southampton, SO9
4XY, UK
REMARK revised by [3]
REFERENCE 2 (bases 43 to 1644)
AUTHORS Watson, M.W., Lambden, P.R., Ward, M.E. and Clarke, I.N.
TITLE Chlamydia trachomatis 60kDa cysteine rich outer membrane
JOURNAL protein:sequence homology between trachoma and LCV biovars
FEMS Microbiol. Lett. 65, 293-297 (1989)
REFERENCE 3 (bases 1 to 2385)
```

```
Watson, M.W.
Direct Submission
Submitted (02-FEB-1993) Watson M.W., Southampton University England,
Department of Microbiology, Faculty of Medicine, South Lab and Path
block, Southampton General Hospital, Tremona road, Southampton, SO9
4XY, UK
Data kindly reviewed (06-SEP-1990) by Watson M.

FEATURES
source Location/Qualifiers
    1..2385
        /organism="Chlamydia trachomatis"
        /strain="B/Jali20/Ot."
        /db_xref="taxon:813"
    RBS
    gene 216..482
        /gene="9kDa Crp"
    CDS 216..482
        /gene="9kDa Crp"
        /codon_start=1
        /transl_table=11
        /protein_id="CAA37587.1"
        /db_xref="GI:40682"
        /translation="MKKTALLAALCSVLSLSCCRIVDCFCDFPCAP10CSPCESKKK
        VDVGNCNCGTVFPACKPCGGDTHQDAKHGPQARGIPVDGRKQ"
    RBS 637..641
    gene 647..2290
        /gene="60kDa Crp"
    CDS 647..2290
        /gene="60kDa Crp"
        /codon_start=1
        /transl_table=11
        /protein_id="CAA37588.1"
        /db_xref="GI:40683"
        /translation="MNKLIRPAVTIFAFTSVASLFGVLETSMAESLSTNVLSLADT
        KAKDNTSHKSKARKNHSHKETPVDRKEVAPVHESKATGPKQDCFCGRMYTVKVNDRN
        VEITQAVPEYATVSGSPYIEITATGKEDVDVITTOOLPCBAEFVRSOPATPTADCK
        LWKIDRLGCOGEKSKITVWPKPLAKGCCFTATVACPEIRSYTVTCGQPAICVQEGP
        ENKLURCPVYIKINIVNGTATARNVVPENPVDGTAHSGGQVLPFTLGDMPQEHRR
        TITVEFCPLKRRATNATVSYCGHKNATASVTVINPEPCQVSTAGADWSVTCKPVE
        VYVSNPGLDLVRDVEDTLSPGTVLEAAGQISCNKRVVTKELNDSVLSLOKYK
        LVRACTPGQFTNNVVKSCDCGCTSCAEATYWKGVAAITHMCVVVDTPCVCGET
        VYRICVTNRGSAEDTNVSLMKFSEKLPVFSFGPTKGTITGNTVVFOSLPRIGSKET
        VEFSVTLKAVSAGDARGEAILLSDTLTPVPSDENTHIY"
BASE COUNT 708 a 440 c 522 g 715 t
ORIGIN

Query Match 42.1% Score 888 DB 1: Length 2385:
Best Local Similarity 70.9% Pred. No. 7.7e-190;
Matches 1248; Conservative 0; Mismatches 480; Indels 32; Gaps 4:

QY 61 aaatgttgagggttaa---gagtttcaaaaacattctaccgatgagcagaagaataaat 116
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 566 AAATGTTGAGGGTAAAAGTTAGTTAATACCAATTTCTACCCGATGGCAGACAAAAATAA 625
QY 117 aaacatcgatagagatcccttatgtcccaaaactcatcagacgagtagttacggtctctgc 176
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 626 TCTATGCGAATAGAGAT-CCTATGAACAAATCATCAGACGACGACGATCTTCGC 684
QY 177 gcaacagagatagcaggttgccttccagcgggggatagagcccgctgtacagagtc 236
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 685 GGTGACTAGTGTGCGAGTTTATTGCTGTCGGGGTGTATAGAGACCTCTATGCCACAGTC 744
QY 237 tctgattactaagatcgtctcgtcagtcggaacaaacccagcacctgttctctatgcagc 296
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 745 TCTCTCTCAACAGTTATTAGCTTAGCTTGACACCAAA-----AGC 783
QY 297 gaagaaggttagacttgcctcgtagaaataaaacacccagttgaaacaaacacccggtgtgc 956
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 784 GAAAGACACACTTCTCATAAAGCAAAAAAGCAAGAAAAAACACCAAGACAGACTCC 843
QY 357 ttttggataaagaattttatccctgtgaagaggagcagatgtcaacacctgtagaggtc 916
```



```
/protein_id="AAA23154.1"
/db_xref="GI:144558"
/translation="MNKLIRRAVTFIAVTSVLSFASGVLETSMAESLSTNVISLADT
KADNTSHKSKARKNHSKETLVDRKEVAPVHESKATGPKQDSFCGRWYTVKVNDRN
VITQAVPEYATVSGSPYIEITATGKRDCVDVLIITQOLPCEAEFVRSOPATPTADGK
LVWKIDRLCOCEKSKITVWPKLKEGCCCTAATVCACPEIRSVTKGQOPATCVKQEGP
ENALRUCPVYKLVNVOGTALARNVVENPVDPGVYHSGGSRVLTFILGDMQGEHR
TITVEFCFLPKGRATNLTATVSYCGGHKNTASVTVINERCVQVSLAGADWSVCKPVE
XVLSVNSQEDLVLRDQVVEDTLSPGTVLEAAGAOISCNKVVWTKELNPFESLQYKV
LVRAATPGQFTNNVVVSKSDCGTCTSCAEATYTWGAYAAATHMVCVDCDPCVCGENT
VYRICVTNKGASDINLSMLKFSKELQVPSFGPKGTHITGNTVVFDSLPLGSKET
VFESVTLKAVSAGARGAEILSDTLTVFVPSVDENHIY".
BASE COUNT      712 a 436 c 525 g      717 t
ORIGIN

Query Match      42.1%; Score 888; DB 1; Length 2390;
Best Local Similarity 70.9%; Pred. No. 7.7e-190;
Matches 1248; Conservative 0; Mismatches 480; Indels 32; Gaps 4:

Qy 61 aaatgttgagggtaa----gagttacaaaacattctaccgcatgagcagaagaaaaaat 116
Db 555 AAATGTTGAGGGTAAAGTTAGTTAATAACAATTCTACCGGATGGCAGACAAAATAA 614
Qy 117 aaacatgcgatagagatccctatgtccaaactcatcaagacagtagttacggtcctgc 176
Db 615 TCTATGCGAATAGGAGAT--CCTATGAACAAACTCATCAGAGGACGATGACGATCTCGC 673
Qy 177 gctaacagatggcgagttgcttccagcgggggtatagagggcgctgtagcagagtc 236
Db 674 GGTGACTAGTGTGGCGAGTTATTTCGTAGCGGGGTGTTAGAGACCTCTATGGCAGAGTC 733
Qy 237 tctgattactaagatcgctgcagtcgcggaacaaacagcacctgttctctatgacagc 296
Db 734 TCTCTCTACAAAGCTATTAGCTTAGCTGACACCAA-----AGC 772
Qy 297 gaagaaggttagactgtccgttagaataaacaaccagttgaacaaaaaacgctggtgc 356
Db 773 GAAGACACACCTTCTATAAAGGCAAAAAGCAAGAAAACCCACAGCAAGAGACTCT 832
Qy 357 tttttgtatagaagaattttatccctgtgaagaggagcagatgtcaacctgtagagctca 416
Db 833 CGTAGACCGTAAAGAGTTGCTCCGGTTCATGAG-----TCTAAAGCTACAGGACTAA 886
Qy 417 gcaagagctctgtacaggaagattgtattctgtaaaagtaaacagatgattgcaacgtaga 476
Db 887 ACAGGATTCCTGTTTGGCAGAAATGTATACAGTCAAAAGTTAATGATCATCGTAATCTGA 946
Qy 477 aatttgccagtcggttccagaacacgactactgtaggattctcttaccctattgaaatcct 536
Db 947 AATCACACAGCTGTTCTCTGAATATGCTACGGTAGGATTCCTCCCTATTCCTAATGAATTC 1006
Qy 537 tgcataaggcaaaaaagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 596
Db 1007 TGCTACAGGTPAAAGGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1066
Qy 597 tgaatttcgaagcagtgatccagagaacaaactcctcacagtgatgggaaattagtcctggaa 656
Db 1067 AGAGTTTCGTACGCAGTGATCCAGCGACAACTCCTACTCTGATGGTAAGCTAGTTTGGAA 1126
Qy 657 aatcgatcgctgggtgcgagagataaatgcaaaattactgtatggttaaaacaccttaa 716
Db 1127 AATTGACCGTTTAGGACAGGCGAAAGAGTAATAATCTGTATGGTAAACCTCTTAA 1186
Qy 717 agaaggttgctgcttcacagctgctactgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 776
Db 1187 AGAAGGTTGCTGTTTACAGCTGCAACAGTATGCGCTGTGTGCGAGATCCGTTCCGGTTAC 1246
Qy 777 taaatgcggttcaaccagccatttattaaagcaagaagagacactgactgtgctgtgcctaag 836
Db 1247 AAAATGTGGACAACCGCTATCTGTGTTAAACAAGAGGCCAGAGAAATGCTTGTGTGCG 1306
Qy 837 atgcccgtgtatgctacaaaaatcgaagtagtgaacacagggatctctgctattcccgtaacgt 896
```

Search completed: May 25, 2002, 19:55:49
Job time: 11616 sec

```
Db 1307 TTCCCCAGTAGTTTACAAAATTAATGTAGTGAACCAAGGAACAGCAATAGCTCGTAACTG 1366
Qy 897 aactgtagataactcctgttcccgatggcgtatctcctacatgcatctgggtcaagaagattctctc 956
Db 1367 TGTGTTGAAATCCCTGTTTCCAGATGGTTACGCTCACTCTTCGGACAGCGTCTACTGAC 1426
Qy 957 ttttaacttagagacagacagacccctggcgataaaaaaggtatttacaggttgagttctgcgc 1016
Db 1427 GTTTACTCTTGGAGATATGCAACCTGGAGAGCACAGAACAAATTAAGTGTAGAGTGTGTGTC 1486
Qy 1017 tcaagaagaaggttcaaaactcaactaaagttgctactgttaacttactcgggtggagacaaaatg 1076
Db 1487 GCTTAAACGTGGTCTGCTTACCAATATAGCAACGCTTCTTACTGTGGAGGACATAAAAA 1546
Qy 1077 tctgcaaaatgtaactacagttgtttaaagccctgttgtaagaagtaaaatctctctggtgc 1136
Db 1547 TACAGCAAGCGTAACAACTGTGATCAATGAGCCTTGGCTACAAGTAAGTATTTCAGGAGC 1606
Qy 1137 tgattggtctctacgtatgtataaacctgtgagctactctctctcagtatcgaatcctcgaga 1196
Db 1607 AGATTGCTCTTATGTTTGTAGCCCTGTAGAATATGTGATCTCCGTTTCCAANTCCTGGAGA 1666
Qy 1197 ctgggttctctatgtagtgcgtgataccaaagatacactccctctctggtgtctacagtaactcga 1256
Db 1667 TCTTGTGTTGGAGATATGCGTCGTTGAAAGACACTCTTCTCCTCCGGAGTCAACATTTCTTA 1726
Qy 1257 agctcctggtggagagatcgtcgttaataaagttgttctggtgattataaagaaalgctgc 1316
Db 1727 AGCTGAGGAGGCTCAAAATTTCTTGTAAATAAAGTAGTTTGGACTGTGAAAGAACTGAATCC 1786
Qy 1317 agagaaaacccctccagtttaaaactgttagtgaagactcaagttcctgtgaagattcacaag 1376
Db 1787 TGGAGAGTCTCTACAGTATAAAGTCTTAGTAGAGACACAACTCCTGGACAATTTACAAA 1846
Qy 1377 tcaagttgagtaactagtagtctaaactcgggaacatgtacatctcttcgacagaaacac 1436
Db 1847 TAATGTTGTTGCAAGAGCTGCTCGACTGTGACTGTGTTGACTTCTTTCGCGAGAGCGAC 1906
Qy 1437 aacacattgtgaaaggtccttgacgctaccctatgtcggtattagacacaaatgatcctat 1496
Db 1907 AACTTACTGTGAAAGGAGTTGCTGCTACTTCATATGTGGCTAGTAGATACTTGTGACGCTGT 1966
Qy 1497 ctgtgtaggagaaaaatactgtctctatcgtctgtgtaactaaacggtgttctctgtgaga 1556
Db 1967 TTGCTGAGGAGAAAAATACTGTTTACCGTATTTGTGTCCCAACACAGAGGTTTCTGCAGAAGA 2026
Qy 1557 tactacgttatctttaactgtgaagttctcaaaaagacttcaagcaaatagcttcttcagg 1616
Db 2027 TACAATGTTTCTTTAATGCTTAAATCTCTTAAGAAGACTGCAACCTGTATCTCTTCTCTGG 2086
Qy 1617 tcaactaaagaagacatttcaggtaataaccgttctgttcttcgaagcctttacctaataactcgg 1676
Db 2087 ACCAACTAAAGGAAGCATTTACAGCAATACAGTAGTATTCGATTCGTTACTAGATTAGG 2146
Qy 1677 tcttaagggaactctglagagtttctctgttacccttgaaaggtattgtctcccgagatgctcg 1736
Db 2147 TCTTAAAGAACTGTAGAGTTTCTGTAAACATTTGAAAGCAGTATACAGCTGGAGATGCTCG 2206
Qy 1737 cggcgaagctattcttctctgtacactgaacttccagctatcagacacagaabaatc 1796
Db 2207 TGGGAAGCGATTCTTCTTCGATACATTCACCTGTTCCAGTTTCTGATACAGAGAAATAC 2266
Qy 1797 ccacgtgtattaaattctaa 1816
Db 2267 ACACATCTATTAACTCTTGA 2286
```


XX WPI: 2000-618918/59.
DR P-PSDB: AAB18820.
XX
XX
PT New polynucleotides encoding a 60kda cysteine-rich membrane protein
PT from Chlamydia, useful as a vaccine for preventing and treating
PT Chlamydia infection in mammals
XX
XX
PS Claim 2; Fig 1; 77pp; English.
XX
CC The present sequence encodes a Chlamydia 60 kDa cysteine-rich membrane
CC protein. The membrane-rich polynucleotide and polypeptide are useful
CC for preventing or treating Chlamydia infections, such as community
CC pneumonia, upper respiratory tract infections, bronchitis and sinusitis.
CC They are also useful for diagnosing Chlamydia infection by assaying a
CC body fluid of a mammal. The polypeptide is useful for vaccine
CC production.
XX
SQ Sequence 2111 BP: 634 A; 417 C; 442 G; 618 T; 0 other;

Query Match 100.0%; Score 2111; DB 21; Length 2111;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ttgatcaggtagtagagagatgaataattctctgactactacctaattcagataataaac 60
Db 1 ttgatcaggtagtagagagatgaataattctctgactactacctaattcagataataaac 60
Qy 61 aaatgttgagggttaagagtttacaataattctaccgtagcgagagaagaaaaataaac 120
Db 61 aaatgttgagggttaagagtttacaataattctaccgtagcgagagaagaaaaataaac 120
Qy 121 atcgatagagagatccctatgctccaaactcatcagacagtagttagcgtccctgcgcta 180
Db 121 atcgatagagagatccctatgctccaaactcatcagacagtagttagcgtccctgcgcta 180
Qy 181 acagatagtcggaggttgccttcccgaggggggtatagagccgctgtagcagagttcttg 240
Db 181 acagatagtcggaggttgccttcccgaggggggtatagagccgctgtagcagagttcttg 240
Qy 241 atactaagatgcctgcctagtcggaacaaagccagcaccctgttctctatgacagcgag 300
Db 241 atactaagatgcctgcctagtcggaacaaagccagcaccctgttctctatgacagcgag 300
Qy 301 aaggttagactgtccgttagaataaacaaccagttgaacaaaagccgtgtgtcttt 360
Db 301 aaggttagactgtccgttagaataaacaaccagttgaacaaaagccgtgtgtcttt 360
Qy 361 tggataaagaattttatccctgtgaagagggaagatgtcaacctgtagaggtcagcaa 420
Db 361 tggataaagaattttatccctgtgaagagggaagatgtcaacctgtagaggtcagcaa 420
Qy 421 gagtctgtcacggaagatgtattctgtbaaagtaaacgatgattgcaacgtagaatt 480
Db 421 gagtctgtcacggaagatgtattctgtbaaagtaaacgatgattgcaacgtagaatt 480
Qy 481 tgcagtcctgtccagaaacacgctactgttaggatctctaccctattgaaactctgtct 540
Db 481 tgcagtcctgtccagaaacacgctactgttaggatctctaccctattgaaactctgtct 540
Qy 541 ataggcaaaaagattgttattgtattgtattgattacacacagctaccctgcgaagctgaa 600
Db 541 ataggcaaaaagattgttattgtattgtattgattacacacagctaccctgcgaagctgaa 600
Qy 601 ttcgtaagcagtgatcccaagaaactcctcaagtgatgggaaattagttctggaaaatc 660
Db 601 ttcgtaagcagtgatcccaagaaactcctcaagtgatgggaaattagttctggaaaatc 660
Qy 661 gatgcctgggtgcaggaagataaaatgcaaaatctactgtatgggtataaacctcttaagaa 720
Db 661 gatgcctgggtgcaggaagataaaatgcaaaatctactgtatgggtataaacctcttaagaa 720

Qy 721 ggttgctgtctcaacagctgctactgtatgtcttgcacagagctcogttctttatactaaa 780
Db 721 ggttgctgtctcaacagctgctactgtatgtcttgcacagagctcogttctttatactaaa 780
Qy 781 tgcggtcaaccagcatttatttaagcaagaagacctgactgtgcttgcctaaagatgc 840
Db 781 tgcggtcaaccagcatttatttaagcaagaagacctgactgtgcttgcctaaagatgc 840
Qy 841 cctgtatgtctacaaatcgaaatagtgaaacacagatctctattgcccgtaacgaact 900
Db 841 cctgtatgtctacaaatcgaaatagtgaaacacagatctctattgcccgtaacgaact 900
Qy 901 gtatataatcctgttcccgatgcttatctcatgcatctggtcgaagagttcttcttttt 960
Db 901 gtatataatcctgttcccgatgcttatctcatgcatctggtcgaagagttcttcttttt 960
Qy 961 aacttaggagacatgagacctgacgataaaaaggtatttacagttgagttctgcccctaa 1020
Db 961 aacttaggagacatgagacctgacgataaaaaggtatttacagttgagttctgcccctaa 1020
Qy 1021 agaagagggtcaaatcaactaactgcttactgtaacttactgctgacacacaaatgtctt 1080
Db 1021 agaagagggtcaaatcaactaactgcttactgtaacttactgctgacacacaaatgtctt 1080
Qy 1081 gcaaatgtactacagttgttaagtgccttgcgtacaaagtaaatatctctggtgctgat 1140
Db 1081 gcaaatgtactacagttgttaagtgccttgcgtacaaagtaaatatctctggtgctgat 1140
Qy 1141 tggcttactgctgtataaacctgtgagtagcttctctcatcagtatcgaaatccttgagacttg 1200
Db 1141 tggcttactgctgtataaacctgtgagtagcttctctcatcagtatcgaaatccttgagacttg 1200
Qy 1201 gttcttcagtagctgtatccaagatcacactcctctctggtttacagtagctcgaagct 1260
Db 1201 gttcttcagtagctgtatccaagatcacactcctctctggtttacagtagctcgaagct 1260
Qy 1261 cctggtggagagatcgtgtataaagtgttggcggtattaaagaaatgtgtccacgaa 1320
Db 1261 cctggtggagagatcgtgtataaagtgttggcggtattaaagaaatgtgtccacgaa 1320
Qy 1321 gaaacctccagtttaaacctgtagtgaaagctcaagttcctctggaagattcacaacaa 1380
Db 1321 gaaacctccagtttaaacctgtagtgaaagctcaagttcctctggaagattcacaacaa 1380
Qy 1381 gtcgagtagtagtgcctgaacgtagtgcacatgtagcttgcgcagaaacacaa 1440
Db 1381 gtcgagtagtagtgcctgaacgtagtgcacatgtagcttgcgcagaaacacaa 1440
Qy 1441 catggaaagggtcttgcagctaccatgctgctatttagacacaaatgactcttct 1500
Db 1441 catggaaagggtcttgcagctaccatgctgctatttagacacaaatgactcttct 1500
Qy 1501 gtagggaaaaactactgtctatcgtagtgaactaacctggttctgtcgaagatct 1560
Db 1501 gtagggaaaaactactgtctatcgtagtgaactaacctggttctgtcgaagatct 1560
Qy 1561 aacgtatcttaactctgaagtctcacaagaacttcagccaaatagcttcttcagttcca 1620
Db 1561 aacgtatcttaactctgaagtctcacaagaacttcagccaaatagcttcttcagttcca 1620
Qy 1621 actaaaggaacatttcaggttaataaccgttttttcagcgttttaacctgaactcgttt 1680
Db 1621 actaaaggaacatttcaggttaataaccgttttttcagcgttttaacctgaactcgttt 1680
Qy 1681 aaggaatctgtgaggtttctctgttaccctgaaaggtattgtctcccgagagatgctcgac 1740
Db 1681 aaggaatctgtgaggtttctctgttaccctgaaaggtattgtctcccgagagatgctcgac 1740
Qy 1741 gaagctattcttctctgtatcacactgacttccacagtagtccagtagtccagaaataccac 1800
Db 1741 gaagctattcttctctgtatcacactgacttccacagtagtccagtagtccagaaataccac 1800
Qy 1801 gtgtattaaatcttaagggaattatctctaaagcagagcgaattcccgctctgtcttaggat 1860

	Query Match	17.2%;	Score 363.6;	DB 22;	Length 731;
	Best Local Similarity	75.8%;	Pred. NO. 4e-88;		
	Matches 450;	Conservative	0;	Mismatches 144;	Indels 0;
	Gaps				
QY	1223	aagatacacctccctctggtgttacagttactcgaagctcctcgtgtgagagatctcgtta	1282		
Db	4	aagacactctttctcccgagtcacagctcttgaagtcgagagatcgaattctctgta	63		
QY	1283	ataaagtgtgttgcggtattaaagaatagtgccagagagaacccctccagttctaaacttg	1342		
Db	64	ataaagtgtgttgactgtgaagaagtaactgaatcctcgtgagagctctcacagatacaaatgc	123		
QY	1343	tagtgaagctcaagttctcgtgaagattcacaaatcaagtgtcagtaactagtgagctta	1402		
Db	124	tagtaagagacaaactcctcgtgacaattcacaaatactgtgttgtagaagagctgctctg	183		
QY	1403	actcgggaacatgtacatcttgcgcagaaacaacacacattggaagaggtcttgcagcta	1462		
Db		attcttgcgcagagcgacacacttactgtgaagagaggtctgctgcta	243		

[illegible]

Tue May 28 08:55:08 2002

XX 03-DEC-1999; 99US-0454684.
PR 19-APR-2000; 2000US-0556877.
PR 20-JUN-2000; 2000US-0598419.
XX (CORI-) CORIXA CORP.
XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;
XX WPI; 2001-374831/39.
XX Chlamydia polypeptides and fusion proteins useful for preventing pelvic
XX inflammatory disease, trachoma, acute respiratory tract infections,
XX atherosclerosis and heart disease -
XX Example 1; Page 141; 295pp; English.
XX The present nucleotide sequence is provided in a specification
XX relating to compounds and methods for the treatment and diagnosis of
XX chlamydial infection. The compounds provided include polypeptides and
XX fusion proteins comprising immunogenic portions of Chlamydia antigens
XX and DNA sequences encoding such polypeptides. They are useful for
XX vaccinating against chlamydial infection, which causes pelvic
XX inflammatory disease, trachoma, acute respiratory tract infections,
XX atherosclerosis and heart disease.
XX Sequence 269 BP; 83 A; 50 C; 60 G; 76 T; 0 other;

Query Match 8.5%; Score 179.4; DB 22; Length 269;
Best Local Similarity 79.2%; Pred. No. 1.6e-38;
Matches 213; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 470 acgtagaatttcgagtcggttcacagatacgtactgtaggatctctaccctattg 529
Dd 1 atgtgaaacacacagcgtgttcataatgtactcaggtaggatctcctatctgtg 60
QY 530 aaatccttgctataggcgaacaaagattgtgtattgtgattcacacagctacct 589
Dd 61 aaattactgtacaggtgaaagattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 120
QY 590 gcgaagctgaattcgtacgagtcagtcacagaaacactcctacagtgatgggaattag 649
Dd 121 gtgaagcagagttcgtacgagtcagtcacagcagcagcagcagcagcagcagcag 180
QY 650 tctggaaatcgatcgctcgtggtgcagagataaatcgaataattactgtatgggtaaac 709
Dd 181 ttggaaatcgacgcttagcagcagcagcagcagcagcagcagcagcagcagcagcag 240
QY 710 ctcttaagaaggttgcgttcacagct 738
Dd 241 ctcttaagaaggttgcgttcacagct 269

RESULT 8
AAQ06256
ID AAQ06256 standard; DNA; 327 BP.
XX
AC AAQ06256;
XX
DT 12-MAR-1991 (first entry)
XX
DE Chlamydia Crp gene incl. detection primer sequences.
XX
KW PCR; DIANA; solid phase medical diagnosis; ss.
XX
OS Chlamydia trachomatis.
XX
FH Key Location/Qualifiers
FT 1..327
FT CDS /tag= a
FT /product=Crp gene prod
FT 1..21
FT misc_RNA

FT FT /*tag= b
FT FT /label= detection primer RIT 23
FT FT 22..42
FT FT /*tag= c
FT FT /label= detection primer RIT 25
FT FT 244..264
FT FT /*tag= d
FT FT /label= detection primer RIR 26
FT FT 298..318
FT FT /*tag= e
FT FT /label= detection primer RIT 24
XX
XX WO9011369-A.
XX 04-OCT-1990.
XX 15-MAR-1990; 90WO-EP00454.
XX 22-MAR-1989; 89GB-0006642.
XX 22-MAR-1989; 89GB-0006641.
XX (CEMU-) CEMU BIOTEKNIK.
XX Uhlen M;
XX WPI; 1990-320269/42.
XX P-PSDB; AAR07399.
XX Amplification of DNA by PCR - using distal DNA on a primer as a
XX handle for attachment to a solid support or a label, used in
XX diagnosis of medical conditions
XX Disclosure; fig 19; 66pp; English.
XX This gene is detected using the primers RIT 23-26 (see features),
XX Detection of immobilised amplified nucleic acids (DIANA), an
XX improved PCR procedure, is utilised. One of the detection primers
XX is used as a handle, with binding affinity for a partner primer
XX attached to a solid support. The primers RIT 23 and -24 are used
XX to amplify a C. trachomatis serovar L2 cell culture using PCR.
XX A 2nd PCR step is then performed using primers RIT 25 and -26.
XX Sequence 327 BP; 95 A; 57 C; 80 G; 95 T; 0 other;

Query Match 8.1%; Score 171.8; DB 11; Length 327;
Best Local Similarity 70.3%; Pred. No. 2e-36;
Matches 230; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
QY 1045 gctactgtactactcgtggtgacacaaatgtctgcaaatgtaactacagttgta 1104
Dd 1 gcaatggttcttactgtggtgacacaaataacacagagggttaacactgtgatca 60
QY 1105 gacgttgtgtacaaagtaatactctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1164
Dd 61 gacgttgtgtacaaagtaatactctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 120
Dd 1165 gactactctatcagtatcgaatcctggtgagctgtgtgtgtgtgtgtgtgtgtgt 1224
QY 121 gaatatgtatcgt 180
Dd 1225 gatacactccttctgt 1284
Dd 181 gacactcttctcgcgagtcacagttcttgagctcagagctcaattcttctgtgt 240
QY 1285 aaagt 1344
Dd 241 aaagt 300
QY 1345 gtgaagctcaggttctcgtggaagattc 1371
Dd 301 gtaagagcacaactcctcgtggaacattc 327

```

RESULT 9
ID AAA50031 standard: DNA; 800 BP.
XX
AC AAA50031;
XX
DT 10-OCT-2000 (first entry)
XX
DE DNA encoding Chlamydia pneumoniae antigen CPN100696 RY-54.
XX
KW CPN100696 RY-55; antigen; infection; diagnosis; therapy; vaccine;
XX sulfur-rich protein; ds.
XX
OS Chlamydia pneumoniae.
XX
FH Key Location/Qualifiers
CDS 101..691
FT /*tag= a
XX
PN WO200039158-A1.
XX
PD 06-JUL-2000.
XX
XX 23-DEC-1999; 99WO-CA01230.
XX
PR 23-DEC-1998; 98US-0113280.
PR 23-DEC-1998; 98US-0113281.
PR 23-DEC-1998; 98US-0113282.
PR 23-DEC-1998; 98US-0113283.
PR 23-DEC-1998; 98US-0113284.
PR 23-DEC-1998; 98US-0113285.
PR 23-DEC-1998; 98US-0113385.
PR 23-DEC-1998; 98US-0114050.
PR 23-DEC-1998; 98US-0114056.
PR 23-DEC-1998; 98US-0114057.
PR 23-DEC-1998; 98US-0114058.
PR 23-DEC-1998; 98US-0114059.
PR 23-DEC-1998; 98US-0114061.
XX
PA (CONN-) CONNAUGHT LAB LTD.
XX
PI Murdin AB, Oomen RP, Wang J;
XX
DR WPI; 2000-452369/39.
XX
PT P-PSDB; AAY95544.
XX
PS Novel Chlamydia polynucleotides and polypeptides useful for diagnosis,
prevention and treatment of Chlamydia infection in mammals -
XX
XX Claim 2(a); Page 58-59; 215pp; English.
XX
XX The present sequence is that of Chlamydia pneumoniae genomic DNA
including an open reading frame that codes for CPN100696 RY-55 (see
AAY95544), a putative sulfur-rich protein. It is an example of
C. pneumoniae polynucleotide molecules of the invention (see
AAA50030-42) that encode antigenic polypeptides (see AAY95543-55) useful
in the diagnosis, treatment and prevention of Chlamydia infection.
XX The polynucleotides can be utilised: in the recombinant production
of Chlamydia antigens using transformed unicellular host cells; in
XX vaccines or live vaccine vectors; in naked form or formulated with
a delivery vehicle for therapy and prophylaxis of Chlamydia
XX infection; and as probes or primers for diagnosis of Chlamydia
XX infection.
XX
SQ Sequence 800 BP; 239 A; 150 C; 162 G; 249 T; 0 other;

Query Match 5.7%; Score 120.8; DB 21; Length 800;
Best Local Similarity 98.4%; Pred. No. 1.8e-22;
Matches 122; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1988 ttattttaaaagcccatcttttaggtatgtaatttaaaatttttaattgaagcttttcta 2047

```

```

Db 1 ttattttaaaagcccatcttttttaggtatgtaatttaaaatttttaattgaagcttttcta 60
QY 2048 gtgtaacctgtcttttttaggaactacactaggagagacgggtatgtcatcaaatctacatcc 2107
Db 61 gtgtaacctgtcttttttaggaactacactaggagagacgggtatgtcatcaaatctacatcc 120
QY 2108 cgta 2111
Db 121 cgta 124

RESULT 10
ID AAA70152 standard; DNA; 6033 BP.
XX
AC AAA70152;
XX
DT 07-NOV-2000 (first entry)
XX
DE Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:285.
XX
KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
XX antimalarial; malaria; protozoacide; infection; insecticide; ds.
XX
OS Plasmodium falciparum.
XX
PN WO200025728-A2.
XX
PD 11-MAY-2000.
XX
PF 05-NOV-1999; 99WO-US26796.
XX
PR 05-NOV-1998; 98US-0107131.
XX
PA (HOFF/) HOFFMAN S.
PA (CARU/) CARUCCI D.
PA (GARD/) GARDNER M.
PA (VENT/) VENTER J C.
XX
PI Hoffman S, Carucci D, Gardner M, Venter JC;
XX
DR WPI; 2000-365347/31.
XX
PT Proteins encoded by chromosome 2 of the human malarial parasite,
Plasmodium falciparum, useful as antimalarial vaccines and in the
diagnosis of P.falciparum infection -
XX
XX Disclosure; Page 493-495; 577pp; English.
XX
XX The present invention describes proteins and their fragments (I) encoded
by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
XX Also described are: (1) nucleotide sequences (II) encoding (I); and (2)
XX vaccines against P. falciparum infection comprising (I) or (II).
XX (I) and (II) are useful for the development of vaccines against
XX P. falciparum infection. (I) and polyclonal antisera or a monoclonal
XX antibody raised to immunogens comprising the sequences of (I), are
XX useful in the detection of infection with P. falciparum. Furthermore,
XX (I) (especially when they are rifins or secreted or membrane proteins)
XX can aid the identification of drugs to treat or prevent P. falciparum
XX infection, or they can be used to identify drug resistance in
XX P. falciparum. Sequencing of the Plasmodium chromosome 2 and the
XX subsequent identification of proteins encoded by it will help to expand
XX our understanding of parasite biology, a process hampered by the
XX complexity of the parasitic lifecycle, and provide new targets for
XX vaccine and drug development. Parasite resistance to drugs and mosquito
XX resistance to insecticides have led to a resurgence of malaria in many
XX parts of the world, and there is a pressing need for vaccines and new
XX drugs. AAA70078 to AAA70287 and AAB1814 to AAB18352 represent nucleotide
XX and protein sequences given in the present invention, but which are not
XX specifically mentioned within the specification.
XX
SQ Sequence 6033 BP; 3019 A; 437 C; 707 G; 1870 T; 0 other;

```

us-09-523-647-1-1.rng

Tue May 28 08:55:08 2002

These peptides also induced ICM, indicating that infection with Chlamydia may be involved in the development of ICM. Inflammatory cardiomyopathy peptides are used to determine the risk of ICM by incubation with a subject's T cells and measuring the degree of proliferation (an increased degree being indicative of risk) or to raise specific antibodies which can be used therapeutically and for the detection of Chlamydia. Such peptides can also be used with an adjuvant and an excipient in a vaccine for decreasing ICM.

Sequence 48 BP; 11 A; 10 C; 14 G; 13 T; 0 other;

Query Match 2.2%; Score 46; DB 20; Length 48;
Best Local Similarity 100.0%; Pred. No. 0.0087;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 ggtagagggccgctgtacagagctctctgattactaagatcgctg 256
|||||
DB 1 ggtagagggccgctgtacagagctctctgattactaagatcgctg 46

RESULT 12
AAZ9167
ID AAZ99167 standard; DNA; 48 BP.
XX
AC AAZ99167;
XX
XX 21-JUN-2000 (first entry)
DT
XX Chlamydia pneumoniae Cys-rich OMP partial coding sequence.
DE
XX Cardiant; murine alpha myosin heavy chain; inflammatory myocarditis;
KW autoimmune inflammatory cardiomyopathy; Chlamydia; antibody; vaccine;
KW hybridization probe; ss.
XX
XX Chlamydia pneumoniae.
OS
XX US6034230-A.
FN
XX 07-MAR-2000.
PD
XX 03-MAY-1999; 99US-0303862.
PF
XX 12-AUG-1998; 98US-0133774.
PR
XX (AMGE-) AMGEN CANADA INC.
PA
XX Neu N, Penninger JM, Bachmaier K, Hessel AJ;
PI
XX WPI: 2000-255712/22.
DR P-PSDB; AAY83817.
XX
PT DNA molecules encoding novel myocardial peptides used for inhibiting
PT and inducing inflammatory cardiomyopathy in vivo
XX
XX Claim 1; Column 19; 17pp; English.

This sequence represents the coding sequence for a fragment of the Chlamydia pneumoniae Cys-rich outer membrane protein (OMP). The sequence was obtained by homology searches of the PIR database using the murine M7A-alpha/beta peptide consensus sequence (Y83813) as the query. The screen isolated the peptides Y83814-Y83819 and their corresponding coding sequences 299164-299169. The invention relates to the isolation of sequences coding for peptide sequences derived from bacteria and viruses which may cause inflammatory cardiomyopathy. The peptide sequences are searched based on the sequence of the M7A peptides derived from the murine alpha myosin heavy chain polypeptide. The peptides encoded by the DNAs are used, alone or in conjunction with other therapeutics, for inducing or inhibiting inflammatory cardiomyopathy in vivo, where the cardiomyopathy is autoimmune inflammatory cardiomyopathy, and inflammatory cardiomyopathy caused by Chlamydia or other bacterial or viral infections that cause inflammatory cardiomyopathy. The peptides may also be used for increasing inflammatory

Query Match 2.3%; Score 48; DB 21; Length 6033;
Best Local Similarity 55.4%; Pred. No. 0.022;
Matches 93; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 1944 ttcttaatacaagatattatttaagttcttgaaatccctaaagatttttttaaaagccca 2003
|||||
DB 4366 TTTCGGCTTTCATATAATTTATTACTATAACAATAAGATTTTTCATTATTCGAA 4307

QY 2004 tcttttagtgatgaataaaatttttaataagctttcttcctagtgtaacctgtcttt 2063
|||||
DB 4306 CATTACATAATTTGTTACTTAATTTTAAATTTTATTTTATTAATGAATTAATCTA 4247

QY 2064 taggaactacactaggaagacggtatgtcatcaaatcatatcccgta 2111
|||||
DB 4246 AAGATATAAAGCTAGCAAAAGGAGGTATTTTCATATCTTTTTCGPA 4199

RESULT 11
AAZ28184
ID AAZ28184 standard; DNA; 48 BP.
XX
AC AAZ28184;
XX
XX 20-DEC-1999 (first entry)
DT
XX Chlamydia pneumoniae outer membrane protein-derived peptide DNA.
DE
XX Heart disease; inflammatory; autoimmune; cardiomyopathy; myosin;
KW Chlamydia; induction; vaccine; ds.
KW
XX Synthetic.
OS
XX Chlamydia pneumoniae.
XX
XX Key Location/Qualifiers
FH 1.48
FT CDS
FT /*tag= a
FT /product= "Chlamydia pneumoniae outer membrane
FT protein-derived peptide"
FT /transl_except= (pos: 46..48, aa:Ala)

US5962636-A.
PN
XX
XX 05-OCT-1999.
PD
XX 12-AUG-1998; 98US-0133774.
PF
XX 12-AUG-1998; 98US-0133774.
PR
XX (AMGE-) AMGEN CANADA INC.
PA
XX Bachmaier K, Hessel AJ, Penninger JM, Neu N;
PI
XX WPI: 1999-589735/50.
DR P-PSDB; AAY42728.
XX
XX Peptides that induce or suppress inflammatory cardiomyopathy
PT
XX
PS Example 1; Column 19; 17pp; English.

This sequence represents DNA encoding Chlamydia pneumoniae outer membrane protein-derived peptide, which induces inflammatory cardiomyopathy (ICM) via an autoimmune response in mice immunised with it. The peptide contains an amino acid sequence motif Maxxs (AAY42722) which appears to be required for the induction of murine alpha myosin heavy chain-derived peptide, M7A-alpha, when it was compared with a peptide derived from a homologous region of the murine beta myosin heavy chain, M7A-beta (AAY42724) which did not cause the disease. Several peptide fragments containing the motif were identified from a database and were found to be fragments of cysteine rich outer membrane proteins from various species of Chlamydia.

CC myocarditis in a mammal. Antibodies against the peptides and the peptides
CC themselves are used for measuring the risk of inflammatory cardiomyopathy
CC in a mammal. The peptides may also be used in vaccines. Nucleic acids
CC encoding the peptides may be used as hybridization probes, e.g. in
CC diagnostic assays to test for the presence of Chlamydia DNA.
XX
SQ Sequence 48 BP; 11 A; 10 C; 14 G; 13 T; 0 other;

Query Match 2.2%; Score 46; DB 21; Length 48;
Best Local Similarity 100.0%; Pred. No. 0.0087;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 211 ggtatagaggcgtgtgacgagctctgtgattactaagatcgtcg 256
Db 1 ggtatagaggcgtgtgacgagctctgtgattactaagatcgtcg 46
|||||

RESULT 13
ABL33396
ID ABL33396 standard; DNA: 5641 BP.
XX
AC ABL33396;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 1369.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.

OS Homo sapiens.
XX
XX WO200200928-A2.
PN
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX

PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
PS Claim 1; SEQ ID NO 1369; 32pp + Sequence Listing; German.
XX

CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX

SQ Sequence 5641 BP; 1442 A; 34 C; 930 G; 3235 T; 0 other;

Query Match 2.1%; Score 43.6; DB 24; Length 5641;
Best Local Similarity 54.4%; Pred. No. 0.49;
Matches 87; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Best Local Similarity 58.5%; Pred. No. 0.33;
Matches 76; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
Qy 1906 ttttttggtttataagctcttcaatcccaatcgtagagttcttcaatcaagatatatt 1965
Db 3011 ttttttggtttataagctcttcaatcccaatcgtagagttcttcaatcaagatatatt 1965
|||||
Qy 1966 taagtcttcgaaatcctaagatttttttaaaagccatcttttaggtatgtaataaa 2025
Db 3071 tttgttagtttaattattatttttagtttgattataaaattttttatttagttattgta 3130
|||||
Qy 2026 atttttaatt 2035
Db 3131 tatttttgatt 3140
|||||

RESULT 14
ABL33952
ID ABL33952 standard; DNA: 7819 BP.
XX
AC ABL33952;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 1925.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.

OS Homo sapiens.
XX
XX WO200200928-A2.
PN
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX

PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
PS Claim 1; SEQ ID NO 1925; 32pp + Sequence Listing; German.
XX

CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX

SQ Sequence 7819 BP; 1808 A; 101 C; 1802 G; 4108 T; 0 other;

Query Match 2.0%; Score 43.2; DB 24; Length 7819;
Best Local Similarity 54.4%; Pred. No. 0.49;
Matches 87; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

[illegible]

RESULT	15
ABL34606	
ID	ABL34606 standard; DNA; 7819 BP.
XX	
XX	
XX	
XX	ABL34606;
XX	
XX	
DT	26-MAR-2002 (first entry)
XX	
XX	
DE	Human metastasis associated gene SEQ ID NO: 159.
XX	
XX	
XX	Metastasis associated gene; cytostatic; gene therapy; cancer;
KW	metastasis methylation; gene; ds.
XX	

DS Homo sapiens.
XX
XX
PN WO20017376-A2.
XX
XX 18-OCT-2001.
PD

XX	
PF	06-APR-2001; 2001WO-EF03970.
XX	
PR	06-APR-2000; 2000DE-1019058.
PR	06-APR-2000; 2000DE-1019173.
PR	30-JUN-2000; 2000DE-1032529.
PR	01-SEP-2000; 2000DE-1043826.
XX	
XX	(PPTG-) EPIGENOMICS AG.
DA	

XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPT: 2002-010922/01.

xx New nucleic acid derived from chemically treated metastasis genes,
pt useful for diagnosis of cancers by analysis of cytosine methylation,
pt also for treatment -

XX
2300 + Sequence Listing; English.

XX The present invention provides a number of human metastasis associated
CC genes which are modified by cytosine methylation. The sequences can be
CC used in the diagnosis and treatment of cancer. The present sequence is
CC one of the genes of the invention.

XX
sequence 7819 BP: 1808 A; 101 C; 1802 G; 4108 T; 0 other;

Query Match	2.08;	Score 43.2;	DB 24;	Length 7819;
Best Local Similarity	54.4%;			
Matches	87;	Conservative		
Indels	0;	No. Mismatches	73;	Gaps 0;

[illegible][illegible]

0v 2026 atttttaattaaagccttttccctagtgtaacctgcttcctta 2065

Search completed: May 25, 2002, 20:48:39
Job time: 13786 sec

QY 533 tccttgctatggcaaaaagattgctgttgattacacacagctacottg 592
Db 61 TCCTTGCTATAGGCAAAAAGATTGTTGATGTGATTACACACACTACCTTGG 120
QY 593 aagctgaattcgaagcagtgatcagaaacaaactcctcacaagtgtggaaattagtct 652
Db 121 AAGCTGAATTCGTAAAGCAGTATCCAGAAACAACTCCTTACAAAGTGATGGAAATAGTCT 180
QY 653 ggaataatcgactgctgggtgagagataaaatgcaaaattactgtatgg 702
Db 181 GGAATAATCGATCGCTGGTGCAGGAGATAAATGCAAAATTACTGTATGG 230

RESULT 2

US-08-477-270-19
; Sequence 19, Application US/08477270
; Patent No. 5629158
; GENERAL INFORMATION:
; APPLICANT: UHLEN, Mathias
; TITLE OF INVENTION: SOLID PHASE DIAGNOSIS OF MEDICAL
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,270
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/261,010
; FILING DATE:
; APPLICATION NUMBER: US 07/781,157
; FILING DATE: 07-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16787/153 DFBC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 327 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Other nucleic acid;
; DESCRIPTION: Synthetic DNA oligonucleotide
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..327
; FEATURE:
; NAME/KEY: misc_difference
; LOCATION: replace(60, "")
; OTHER INFORMATION: /note= "The C at position 60 can
; OTHER INFORMATION: also be a T. The amino acid at position 20 will
; OTHER INFORMATION: be Asn in either case."
; FEATURE:
; NAME/KEY: misc_difference
; LOCATION: replace(178, "")
; OTHER INFORMATION: /note= "The A at position 178 can
; OTHER INFORMATION: also be a G. If it is a G, then the amino acid at

; OTHER INFORMATION: position 60 Glu instead of Lys."
US-08-477-270-19

Query Match 8.2%; Score 173.4; DB 1; Length 327;
Best Local Similarity 70.6%; Pred. No. 2.6e-41;
Matches 231; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
QY 1045 gctactgaacttactcggtggacacaaaatttctgcacaaatgtacactacagttgtta 1104
Db 1 GCATGTTCTTCTTACTGTGGAGGACATAAAATAATACAGCAAGGTAACAACTGTGATCAAC 60
QY 1105 gagccttggtacaaataatctctggtgctgattggtcttactgattgtaaaacctgtg 1164
Db 61 GAGCCTTTCGCTACAAAGTAAGTATTGCGAGGACAGATTTGCTTATGTTTGAAGCCTGTA 120
QY 1165 gagtactctatctcagtatcgaatcctggagacttggtttcttctatgatgctgatac 1224
Db 121 GAATATGTGATCTCCGTTTCCAATCTGTGGAGATCTTGTGTCCGAGATGTCGTCGTTAAA 180
QY 1225 gatactcctctctggtgttacagtaactcgaagctcctggtgagagatctgctgta 1284
Db 181 GACACTCTTTCTCCGGAGTCACAGTTCTTTGAAGCTGCAGAGCTCAAAATTTCTTGTAA 240
QY 1285 aaagttgttgcggtattaaagaaatggtgccagaggagaaacccctccagtttaaa 1344
Db 241 AAAGTAGTTGGACTGTGAAAGAACTGAATCTCTGGAGAGTCTCTACAGTATAAAGTTCTA 300
QY 1345 gtgaaagctcaagttcctggaagattc 1371
Db 301 GTAAGAGACAAACTCCTGGACAAATTC 327

RESULT 3

US-09-133-774-20
; Sequence 20, Application US/09133774B
; Patent No. 5962636
; GENERAL INFORMATION:
; APPLICANT: Bachmaier, Kurt
; APPLICANT: Hessel, Andrew J.
; APPLICANT: Neu M.D., Nikolaus
; APPLICANT: Penninger, Josef M.
; TITLE OF INVENTION: No. 5962636e1 Peptides Capable of Modulating Inflammatory Hear
; FILE REFERENCE: A-536
; CURRENT APPLICATION NUMBER: US/09/133,774B
; CURRENT FILING DATE: 1998-08-12
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 48
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-133-774-20

Query Match 2.2%; Score 46; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 211 ggtatagagccgctgtagcagagctctctgattactaagatcgctg 256
Db 1 ggtatagagccgctgtagcagagctctctgattactaagatcgctg 46

RESULT 4

US-09-303-862-20
; Sequence 20, Application US/09303862
; Patent No. 6034230
; GENERAL INFORMATION:
; APPLICANT: Bachmaier, Kurt
; APPLICANT: Hessel, Andrew J.
; APPLICANT: Neu M.D., Nikolaus

ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 222/005
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 1508 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-714-918-31

Query Match 1.9%; Score 39.6; DB 3; Length 1508;
Best Local Similarity 49.5%; Pred. No. 0.08;
Matches 102; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 503 ctactgtaggatctctaccattgaaatccttctctataggaacaaagattgtgttg 562
Db 1268 CAATTGAGATTTCATTTAAATAAGAAATTTTCGAAAAAGAAATTAGAGGCTTTA 1209
QY 563 atgttgtgattacacacagctacctcggaagctgaattctgaagcagdgatccagaaa 622
Db 1208 ATATCACATTCTCATAAAGAACGTATCATACCGTATTTAGATTATGTTGATGAACAAG 1149
QY 623 caactcctacaagtgaggaaattagtcctggaaatcgatcgctgggtgcagagata 682
Db 1148 CGATTAAATGCAGGTGCAGTTAACACTGTTTGTGATAAAGATGCGCAAGTGGTAGGTATA 1089
QY 683 aatgcaaaattactgtatgggtataa 708
Db 1088 ATACAGATGGTATGGTTATGTTAAA 1063

RESULT 7

US-09-265-315-31/C
Sequence 31, Application US/09265315
Patent No. 6187541
GENERAL INFORMATION:
APPLICANT: Benton, Bret
APPLICANT: Lee, Ving J.
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmid, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
FILING DATE: March 9, 1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/714,918

FILING DATE: September 13, 1996
APPLICATION NUMBER: 60/009,102
FILING DATE: December 22, 1995
APPLICATION NUMBER: 60/003,798
FILING DATE: September 15, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 240/247
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 1508 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-265-315-31

Query Match 1.9%; Score 39.6; DB 4; Length 1508;
Best Local Similarity 49.5%; Pred. No. 0.08;
Matches 102; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 503 ctactgtaggatctctaccattgaaatccttctctataggaacaaagattgtgttg 562
Db 1268 CAATTGAGATTTCATTTAAATAAGAAATTTTCGAAAAAGAAATTAGAGGCTTTA 1209
QY 563 atgttgtgattacacacagctacctcggaagctgaattctgaagcagdgatccagaaa 622
Db 1208 ATATCACATTCTCATAAAGAACGTATCATACCGTATTTAGATTATGTTGATGAACAAG 1149
QY 623 caactcctacaagtgaggaaattagtcctggaaatcgatcgctgggtgcagagata 682
Db 1148 CGATTAAATGCAGGTGCAGTTAACACTGTTTGTGATAAAGATGCGCAAGTGGTAGGTATA 1089
QY 683 aatgcaaaattactgtatgggtataa 708
Db 1088 ATACAGATGGTATGGTTATGTTAAA 1063

RESULT 8

US-09-265-315-31/C
Sequence 31, Application US/09265315
Patent No. 6187541
GENERAL INFORMATION:
APPLICANT: Benton, Bret
APPLICANT: Lee, Ving J.
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmid, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:

Tue May 28 08:55:08 2002

APPLICATION NUMBER: US/09/265,315
FILING DATE: March 9, 1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/714,918
FILING DATE: September 13, 1996
APPLICATION NUMBER: 60/009,102
FILING DATE: December 22, 1995
APPLICATION NUMBER: 60/003,798
FILING DATE: September 15, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 240/247
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 1508 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-265-315-31

Query Match 1.9%; Score 39.6; DB 4; Length 1508;
Best Local Similarity 49.5%; Pred. No. 0.08;
Matches 102; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 503 ctactgtaggatctctaccctattgaaatccttgctataggcgaagaaagattgtgtg 562
DB 1268 CAATTGAAGATTTTCATTTAAATAAGAAATTTTCGAAAAAAGAAATTTAGAAAGCTTTA 1209

QY 563 atgtgtgattacacacagctaccttcgaaagctgaattcgaagcagctgacccagaaa 622
DB 1208 ATATCACAATTCCTCATAAAGAACGTAACACCGTATTAGATTATGTTGATGAACAAG 1149

QY 623 caactcctacaagtgatggaaattagctcgaaataatcgcctcggtgcaggagata 682
DB 1148 CGATTAAATGCAGTGCAGTTAACACTGTTTGTATAAAGATGCAAGTGGATAGGTTATA 1089

QY 683 aatgcaaaattactgtatgggtaaaa 708
DB 1088 ATACAGATGTTATGTTATGTTATAA 1063

RESULT 9
US-09-266-417-31/C
Sequence 31, Application US/09266417
Patent No. 6228586
GENERAL INFORMATION:
APPLICANT: Benton, Bret
APPLICANT: Lee, Ving J.
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmid, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
TITLE OF INVENTION: TARGET GENES
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/266,417
FILING DATE: March 9, 1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/714,918
FILING DATE: September 13, 1996
APPLICATION NUMBER: 60/009,102
FILING DATE: December 22, 1995
APPLICATION NUMBER: 60/003,798
FILING DATE: September 15, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 240/248
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 1508 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-266-417-31

Query Match 1.9%; Score 39.6; DB 4; Length 1508;
Best Local Similarity 49.5%; Pred. No. 0.08;
Matches 102; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 503 ctactgtaggatctctaccctattgaaatccttgctataggcgaagaaagattgtgtg 562
DB 1268 CAATTGAAGATTTTCATTTAAATAAGAAATTTTCGAAAAAAGAAATTTAGAAAGCTTTA 1209

QY 563 atgtgtgattacacacagctaccttcgaaagctgaattcgaagcagctgacccagaaa 622
DB 1208 ATATCACAATTCCTCATAAAGAACGTAACACCGTATTAGATTATGTTGATGAACAAG 1149

QY 623 caactcctacaagtgatggaaattagctcgaaataatcgcctcggtgcaggagata 682
DB 1148 CGATTAAATGCAGTGCAGTTAACACTGTTTGTATAAAGATGCAAGTGGATAGGTTATA 1089

QY 683 aatgcaaaattactgtatgggtaaaa 708
DB 1088 ATACAGATGTTATGTTATGTTATAA 1063

RESULT 10
US-08-938-546-5/C
Sequence 5, Application US/08938546
Patent No. 6004556
GENERAL INFORMATION:
APPLICANT: Fosberry, Andrew P.
APPLICANT: Lawlor, Elizabeth J.
APPLICANT: Nicholas, Richard O.
TITLE OF INVENTION: No. 6004556el rsbu-1
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

```

RESULT 11
US-09-340-812-5/c
; Sequence 5, Application US/09340812
; Patent No. 6111078
; GENERAL INFORMATION:
; APPLICANT: Fosberry, Andrew P.
; APPLICANT: Lawlor, Elizabeth J.
; APPLICANT: Nicholas, Richard O.
; TITLE OF INVENTION: NO. 6111078el rsbu-1
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/340,812
; FILING DATE:
; CLASSIFICATION:

```

```

RESULT 12
US-09-009-913-1/c
; Sequence 1, Application US/09009913
; Patent No. 6087485
; GENERAL INFORMATION:
; APPLICANT: AXYS Pharmaceuticals, Inc.
; TITLE OF INVENTION: Asthma Related Genes
; NUMBER OF SEQUENCES: 339
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Ave, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009,913
; FILING DATE: 21-JAN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: SEQ-4P
; TELECOMMUNICATION INFORMATION:

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK nucleic - nucleic search, using sw model

Run on: May 25, 2002, 16:51:56 ; Search time 2937.23 Seconds
(without alignments)
9700.320 Million cell updates/sec

Title: US-09-523-647-1
Perfect score: 2111
Sequence: 1 ttgtacagtagtaggaga.....catcaaatctacatccgta 2111

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	57.2	2.7	1101	12	CNS0039G
2	51.4	2.4	1101	12	CNS0039G
3	51.2	2.4	1173	12	CNS06LVV
4	50.8	2.4	928	12	CNS00DKY
5	50.2	2.4	1201	12	CNS016E1
6	48.6	2.3	1101	12	CNS016DT
7	47	2.2	586	12	AZ323116
8	46.6	2.2	698	12	CNS010V2
9	45.8	2.2	976	12	CNS04F5M
10	45.2	2.1	610	9	AA675942
11	45.2	2.1	686	9	AL514705
12	45.2	2.1	822	12	CNS009CW
13	45.2	2.1	987	12	CNS014PO
14	45	2.1	1101	12	CNS001FB
15	44.4	2.1	1205	12	CNS0165A
16	44	2.1	801	12	BH499269
17	43.8	2.1	843	12	CNS00C51

18	43.8	2.1	987	12	CNS014PO
19	43.6	2.1	975	12	CNS0102X
20	43.4	2.1	1225	12	CNS0161D
21	43.2	2.0	945	12	CNS013W4
22	43	2.0	380	9	AW240912
23	43	2.0	843	12	CNS00C51
24	43	2.0	857	12	BH552085
25	43	2.0	1001	12	CNS0155H
26	42.8	2.0	658	9	BA438151
27	42.8	2.0	924	12	CNS07A5L
28	42.8	2.0	1135	12	CNS033GQ
29	42.6	2.0	1101	12	CNS00DT7
30	42.6	2.0	1101	12	CNS017KT
31	42.6	2.0	1102	12	CNS014P4
32	42.4	2.0	802	12	CNS0383B
33	42.4	2.0	1101	12	CNS002AL
34	42.2	2.0	623	12	AQ288608
35	42.2	2.0	945	12	CNS04D0K
36	42.2	2.0	1092	12	CNS020K7
37	42.2	2.0	1168	12	CNS06HDF
38	42	2.0	531	10	BG602527
39	42	2.0	1101	12	CNS00LT2
40	41.8	2.0	298	12	B22209
41	41.8	2.0	367	12	CNS00046
42	41.8	2.0	472	10	BG226909
43	41.8	2.0	578	12	AQ010923
44	41.8	2.0	961	12	CNS008H1
45	41.8	2.0	1061	12	CNS015LM

ALIGNMENTS

RESULT 1
CNS0039G/C
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC # BAC08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION
AL063921
VERSION
AL063921.1 GI:4941778
KEYWORDS
GSS.
SOURCE
fruit fly.
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1101)
Genoscope.
AUTHORS
Direct Submission
TITLE
Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammosier in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"

```

/clone="BACR08K10"
/notes="end : TET3"
BASE COUNT      201 a      64 C      131 g      202 t      503 others
ORIGIN

Query Match      2.7%; Score 57.2; DB 12; Length 1101;
Best Local Similarity 17.2%; Pred. No. 0.0018;
Matches 123; Conservative 275; Mismatches 318; Indels 0; Gaps 0;

QY 1348 aaagctcaagcttcctggaagattcacaaatcaagctgcagtaagctagctgaactgc 1407
Db 1081 DWTGTTTWWKDRDRRWAGDADRWAGDGGAGTWTATWWWWATATWDTWWDKWWWA 1022

QY 1408 ggaacatgcatcttcgcgagaaacacacacatcggaaagctgagctgagctaccatc 1467
Db 1021 TAAKTDTAWTWRWTAWRWAGRGAGKRDRAATDAGAGRGGRKKRDKKDRKXGDG 962

QY 1468 atgctgatttagacacaaatgctctctgtgtgaggagaaatactgtctatcgtatc 1527
Db 961 DDKGKKKKAARAAKAWKWKWDWMDWKWDGAKDRKADDDGAGDKDDGKGDAD 902

QY 1528 tgtgttaactaacctggtctgctgagataactaacgtatctttaaacttcgaagtctoa 1587
Db 901 DDTGDKDDDKDKWDNDKAKGTGDAWAAATDWWMGWADADWMTWDAADADWAD 842

QY 1588 aaagaacttcagcaaatgctcttcaggtccaaactaaaggaacgattcaggaataacc 1647
Db 841 DRWDAAWWDWDAWAGARTADRRDWRGDRAGKRGARKRRDRKADDDKRAADDDDAAT 782

QY 1648 gttgttttcagcgttttactaacctcggttcttaaggaatctgtagagtttctgtacc 1707
Db 781 WTTTWTTRDTDDKWKTDWTRWAADRTWDRDDDDDRAGTAGRKWRRTWKRRKRR 722

QY 1708 tgaagagattgtcccgagagctgcgcggcgaagctattcttcttctctgacacatg 1767
Db 721 DTRWDADADDTRDDRRRRGGDGDAGKGTGKRRRRRDRATWDRDADWADAAWWT 662

QY 1768 actcaccagttatcagacagaaaaatccacgcggtattaaattcgaaggaattatcct 1827
Db 661 TDTDTDDDKDRRRKRGARRRRRTARAADWWTKAWDKAKWKTTRADRWDRWAADTW 602

QY 1828 aaagcagagcagatcttcgctctctgtaggagcttccaaagaagcagcgtcttagta 1887
Db 601 TDARKADRWAKARAWRRDRARAARDRTTKGTTTATTTTAAARAAWAAWAAWAT 542

QY 1888 cctcagtaactaagcggtttttgtttataagctctcctcaatcgaatcgaagttctc 1947
Db 541 TTATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 482

QY 1948 ttaatcaagatatatttaagttctgaaatcctcgaagattatttttaaaagccatctt 2007
Db 481 TTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 422

QY 2008 tttaggtatgaataaaatttttaagcttttctcctagtagtaacctgtctt 2063
Db 421 WTTTWTAAAAAANWWTWTATATKCCCCCCTCCCTCCCTCCCTCCCTCCCTCCCT 366

RESULT 2
LOCUS      CNS0039G
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR08K10 of RPCL-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION  AL063921
VERSION     GSS:
KEYWORDS   fruit fly.
SOURCE     Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

```

```

REFERENCE      1 (bases 1 to 1101)
AUTHORS       Genoscope.
TITLE         Direct Submission
JOURNAL
COMMENT       Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
              BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
              - Web : www.genoscope.cns.fr)
              Determination of this BAC-end sequence was carried out as part of a
              collaboration with the Berkeley Drosophila Genome Project (BDGP).
              The BDGP is constructing a physical map of the Drosophila
              melanogaster genome using these BACs. For further information
              please see http://www.fruitfly.org The BDGP Drosophila
              melanogaster BAC library was prepared by Kazuhiro Oseogawa and
              Aaron Mammeter in Pletier de Jong's laboratory in the Department of
              Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
              NY. The library is named RPCL-98 and was constructed by partial
              EcoRI digestion of Drosophila DNA provided by the BDGP from the
              isogenic strain Y2: cn bw sp, the same strain used for the BDGP's
              PI and EST libraries. A more detailed description of the library
              and how to order individual BAC clones, the entire library, or
              filters for hybridization from the BACPAC Resource Center can be
              found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES       Location/Qualifiers
                1..1101
                /organism="Drosophila melanogaster"
                /db_xref="taxon:7227"
                /clone_lib="RPCL-98"
                /clone="BACR08K10"
                /note="end : TET3"
BASE COUNT    201 a      64 C      131 g      202 t      503 others
ORIGIN

Query Match    2.4%; Score 51.4; DB 12; Length 1101;
Best Local Similarity 14.2%; Pred. No. 0.049;
Matches 84; Conservative 263; Mismatches 244; Indels 0; Gaps 0;

QY 1475 tattagacacaaatgactatctgtgtgaggagaaatactgtctatcgtatctgttaa 1534
Db 510 WTTTAAWAAAAAANAANAANAATAAATWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 11
QY 1535 ctacccgtggttctgctgaagataactaacgtatctttaaacttcgaagcttcaagaac 1594
Db 570 WYUHTTWTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTYTY 11
QY 1595 ttacgacaaatagctcttcaggtccaaactaaaggaacgattcgaagtaacggttctt 1654
Db 630 WHWTYTAAYYYTYCYMYYYHYMHWHHAHAHAHAHAHAHAHAHAHAHAHAHAHAHA 689
QY 1655 tcgacgctttacacaaactcgggttcttaaggaatctgtagagctttctgtaccctgaaag 1714
Db 690 MMCCTTCTCHCYYYHYHTAHHTHTHHWYAHYMYMYMYMYMYMYMYMYMYMYMYMYMY 749
QY 1715 gtattgctccgagagctgcgcgcgaagctatcttcttctgtacactgacttccac 1774
Db 750 HWYHTTWYAWAHAMWMMHHHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHA 809
QY 1775 cagatcacagacagaaaaatccacacgtgtattaaatctcaaggaattcctcaaacag 1834
Db 810 MCTYHCWHYHTAYTCWTWTHHMMWTWTHWYHTTWHHTTTHAAWHHTTTCMMWHHTT 869
QY 1835 agcgatattccgctctgcttagagatagcttccaaagaagctacgcttttagtacctaac 1894
Db 870 WTWATHCACATCTMHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 929
QY 1895 tactaaagcggtttttttgtttttaaagctcttcaatcgaatcgaagcttcttataatca 1954
Db 930 MHHHHHHHHHHHMAWMTWTMTTMMMMSCMHHHHHHHHHHHHHHHHHHHHHHHHHHHH 989
QY 1955 aagattatttaagttctgaaatcctaagatttattttaaaagccatcttctttaggt 2014
Db 990 YUUMCTCHYCTHTWTYTAUWAAWTAHAMTTATWMMHHHAHAHAHAHAHAHAHAHAHA 1049
QY 2015 atgtaataaaatttttaataagcttttctcctagtgtaaacctgcttcttta 2065

```



```

Insert Length: 10000   Std Error: 0.00
Plate: 0044   row: G   column: 06
Seq primer: CGTTGTAACACACGGCCAGT
Class: plasmid ends
High quality sequence stop: 586.

```

FEATURES

1. 586
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG1M004G06"
/clone_lib="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pWD42ny; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (q14732114(gb)/AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent *E. coli* XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN

Query Match	2.2%	Score 47	DB 12	Length 586;
Best local Similarity	52.3%	Prod. No. 0.56;		
Matches 104;	Conservative 0;	Mismatches 95;	Indels 0;	Gaps 0;
DY 1853	tttgagatagcttccaaagaagtagccgcttttagtaacctcaactactaaagcggtttttt	1912		
Db 257	TTTTGGAATATTTATAAAAACTGAATCATCTCCAAATATTAAGTATTGAAATATTTTTCT	198		
DY 1913	gttttaaaagctcttccaatccgaactgtagagttcttcaatcaagaatatatttaagttt	1972		
Db 197	TTTTTAAAAATTCATGGATTTTACCATTAGGTGTTTAAATAATCCAGCTATTAAATTTATTT	138		
DY 1973	ctgaaactccaagatatttttaaagcccatcttttttaggtatgtaactaaagaattttta	2032		
Db 137	GAATATATTAAAGAATATAAATTTTCTTGTAGAAATTAACATAAATACTTTTAAA	78		
DY 2033	attaagccttttccagtgt	2051		
Db 77	ATAAATCTATGCTCCAGT	59		

RESULT 8
CNS010V2/c

CN5010V27C	CNS010V2	698 bp	DNA	linear	GSS 26-JUL-1999
LOCUS					
DEFINITION	Drosophila melanogaster genome survey sequence 17 end of BAC HACN5C08 of DrosEAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL039464				
VERSION	AL039464.1				
KEYWORDS	GSS.				
SOURCE	fruit fly.				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Neopterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 698)				

AUTHORS
TITLE
JOURNAL

AUTHORS	TITLE	JOURNAL	COMMENT
Genoscope.	Direct Submission	Submitted (23-JUL-1999)	Genoscope - Centre National de Sequencage : BP 191006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
			Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome project (BDGP) http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBeloBAC11.

COMMENT

FEATURES SOURCE

```

1. .090
source
/organism="Drosophila melanogaster"
/plasmid="pbelOBAC11"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN05C08"
/notes="end : 117"

```

BASE COUNT	256 a	88 c	42 g	127 t	185 others
ORIGIN	/note- end : 1/				

Query Match	2.2%;	Score 46.6;	DB 12;	Length 698;
Best Local Similarity	35.4%;	Pred. No. 0.73;		
Matches 85;	Conservative 44;	Mismatches 111;	Indels 0	

ov 1794 taccacggtgtattaaattcttaaggaattatccctaaagcagagcgatatattccgctctact 1853

282 KANAMTTKTSKKKTTTWTTTTSTKTAADATTAAATGGAAATGSTASASTGAAAANGNTT 223

0v 1854 ttaggatagctttcagaagaataccgctttagctaccttaccctaaacccgtttttttt 1913

DB 222 KTTCTNTCTNAGSAAATAKAAATKTTTTCMTTAAATKCTTAAATKACCTWAZTTTWTTCCTTCTT 163

[illegible]

163 KTHWTTAHTWTTTAAHTTYTHTTTTNNZACHTTETAAKDZTKSCATDZTAWDZAAATWTZAKTCZAZKSTTA 103

Year	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100																																																																																																																																																		
Population	1071	1072	1073	1074	1075	1076	1077	1078	1079	1080	1081	1082	1083	1084	1085	1086	1087	1088	1089	1090	1091	1092	1093	1094	1095	1096	1097	1098	1099	1100	1101	1102	1103	1104	1105	1106	1107	1108	1109	1110	1111	1112	1113	1114	1115	1116	1117	1118	1119	1120	1121	1122	1123	1124	1125	1126	1127	1128	1129	1130	1131	1132	1133	1134	1135	1136	1137	1138	1139	1140	1141	1142	1143	1144	1145	1146	1147	1148	1149	1150	1151	1152	1153	1154	1155	1156	1157	1158	1159	1160	1161	1162	1163	1164	1165	1166	1167	1168	1169	1170	1171	1172	1173	1174	1175	1176	1177	1178	1179	1180	1181	1182	1183	1184	1185	1186	1187	1188	1189	1190	1191	1192	1193	1194	1195	1196	1197	1198	1199	1200	1201	1202	1203	1204	1205	1206	1207	1208	1209	1210	1211	1212	1213	1214	1215	1216	1217	1218	1219	1220	1221	1222	1223	1224	1225	1226	1227	1228	1229	1230	1231	1232	1233	1234	1235	1236	1237	1238	1239	1240	1241	1242	1243	1244	1245	1246	1247	1248	1249	1250	1251	1252	1253	1254	1255	1256	1257	1258	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275	1276	1277	1278	1279	1280	1281	1282	1283	1284	1285	1286	1287	1288	1289	1290	1291	1292	1293	1294	1295	1296	1297	1298	1299	1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1345	1346

103	102	101	100	99	98	97	96	95	94	93	92	91	90	89	88	87	86	85	84	83	82	81	80	79	78	77	76	75	74	73	72	71	70	69	68	67	66	65	64	63	62	61	60	59	58	57	56	55	54	53	52	51	50	49	48	47	46	45	44	43	42	41	40	39	38	37	36	35	34	33	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
103	102	101	100	99	98	97	96	95	94	93	92	91	90	89	88	87	86	85	84	83	82	81	80	79	78	77	76	75	74	73	72	71	70	69	68	67	66	65	64	63	62	61	60	59	58	57	56	55	54	53	52	51	50	49	48	47	46	45	44	43	42	41	40	39	38	37	36	35	34	33	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1

RESULT 9

LOCUS	CNS04E5M	976 bp	DNA	linear	GSS 21-MAY-2003

103P02 of library G from Tetraodon nigroviridis, genomic survey

AL286627

KEYWORDS GSS; genome survey sequence.

ORGANISM Tetraodon nigroviridis

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei

Tetraodontidae; Tetraodon.

AUTHORS
Roest-Crolius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C.

BOUCHAUD, E.; DATTAG, R.; QUÉLIN, R.; SAGHIAN, R. and
Weissenbach, J.

freshwater pufferfish Tetraodon nigroviridis

JOURNAL OF
COMPUTATIONAL
REFERENCES
2 (bases 1 to 976)

AUTHORS
ROEST-LOUIS, H.; JAILLON, O.; DASTIVA, C.; BOUTEAU, L.; FISHER, S.
Bernot, A.; Fitzames, C.; Wincker, P.; Brattier, P.; Ouetier, F.

TITLE Human gene number estimated by genome wide analysis using
Saurin, W. and Weissbach, J.

Tetraodon nigroviridis DNA sequence
Unpublished
3 (bases 1 to 976)
Genoscope.
Direct Submission
TITLE
JOURNAL
COMMENT
This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES
source
1..976
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="103P02"
/clone_lib="G"
/note="Genoscope sequence ID : COBG103DH01LPL1-end : T7"
BASE COUNT 243 a 151 c 137 g 327 t 118 others
ORIGIN

Query Match 2.2%; Score 45.8; DB 12; Length 976;
Best Local Similarity 36.5%; Pred. No. 1.2;
Matches 113; Conservative 51; Mismatches 145; Indels 1; Gaps 1;
Qy 1792 aataccacggtgtattaaattcgaattatctcctaagcagcgagctattccgctctg 1851
Db 527 WRGAGACGACGACGAGAGTGGTGAANTCACCCTVAGARMGCTTCCCTCGTCTC 586
Qy 1852 cttaagatagattcaagaagctacgcttttagctactactaagcggcttttt 1911
Db 587 CCCTCSTVCBCIVBAVARVYCCBYMMWMMWMMWMMWMMWMMWMMWMMWMMW 646
Qy 1912 tgttttaagctctcaatccatcgtcagagttcttcttaacaaagatatcttaagtt 1971
Db 647 WMTTWT 706
Qy 1972 tctgaatcctcaagatttttttaaaagccatcttttttaggtatgtaattaaattttt 2031
Db 707 T-TAWNTWAT 765
Qy 2032 aattaagctttcttagtgtaacctgtcttttaggaactacacaggagaaagctatgt 2091
Db 766 TTWTWATNATWMTWATWMTWATWMTWATWMTWATWMTWATWMTWATWMTWATW 825
Qy 2092 catcaaatct 2101
Db 826 TTATAAATTT 835

RESULT 10
AA675942
LOCUS
DEFINITION
AA675942
EST
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AA675942 610 bp mRNA linear EST 01-DEC-1997
SMOVL3CAN02G04SK Onchocerca volvulus infective larva cDNA
(SAB94WL-OVL3) Onchocerca volvulus cDNA clone SMOVL3CAN02G04 5',
mRNA sequence.
AA675942
EST
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AA675942.1 GI:2655606
Onchocerca volvulus.
Onchocerca volvulus
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
Onchocercidae; Onchocerca.
1 (bases 1 to 610)
Williams, S.A., Lu, W., Lizotte-Waniewski, M. and Laney, S.J.
Genes expressed in infective third stage larvae of Onchocerca
volvulus
Unpublished (1995)
Contact: Steven A. Williams
Molecular Parasitology
Smith College Department of Biological Sciences
Department of Biological Sciences, Clark Science Center, Smith
College, Northampton, MA, 01063, USA

Tel: 4135853826
Fax: 4135853786
Email: genomesmith.edu
Seq primer: phuescript SK
Location/Qualifiers
1..610
/organism="Onchocerca volvulus"
/strain="Sierra Leone"
/db_xref="taxon:6282"
/clone="SMOVL3CAN02G04"
/clone_lib="Onchocerca volvulus infective larva cDNA
(SAB94WL-OVL3)"
/lab_host="XLI-Blue MRF"
/note="Vector: lambda Unizap XR; Site_1: EcoR I; Site_2:
Xho I; Cutaneous filarial nematode parasite of humans.
mRNA was prepared from third stage infective larvae of
Onchocerca volvulus isolated from mosquitoes 10 days after
infection and converted to double stranded cDNA using
reverse transcriptase and oligo(dT) followed by RNase H
and DNase I. The library had 1.8 x 10⁵ independent
recombinants and average insert size was 900 base pairs.
The library was constructed by Wenhong Lu. The library is
available from Dr. S.A. Williams, email genomesmith.edu."

BASE COUNT 174 a 94 c 72 g 267 t
ORIGIN

Query Match 2.1%; Score 45.2; DB 9; Length 610;
Best Local Similarity 58.0%; Pred. No. 1.6;
Matches 80; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
Qy 1906 ttttttggttttaagctctcctcaacacgtagagttcttcttaacaaagattattt 1965
Db 164 TTTTATTTTACCAATATATATTAAGCAATATTCCTCAATTTCTTTATCAACATCTGCT 223
Qy 1966 taagttcttgaaatcctcaagatttttttaaaagccatcttttttaggtatgtaattaaa 2025
Db 224 AAACATTTTCTTATCAACATCTTTAGTAACATTTCCATCTCTTCTACTGATACATTGAG 283
Qy 2026 atttttaatttaagctttt 2043
Db 284 ATCTTTAACTTTTCTTTTGT 301

RESULT 11
AL514705/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL514705 686 bp mRNA linear EST 13-FEB-2001
AL514705 LTI_NFL006_PL2 Homo sapiens cDNA clone CLOBB013ZF06 3
prime, mRNA sequence.
AL514705
EST
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL514705.1 GI:12778198
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 686)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segreg@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1..686
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CLOBB013ZF06"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end

Tue May 28 08:55:16 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: May 25, 2002, 20:39:16 ; Search time 119.95 Seconds
(without alignments)
514.856 Million cell updates/sec

Title: US-09-523-647-2
Perfect score: 2951
Sequence: 1 MSKLRRVTVLALTSWASC.....ILSSDTLTSPVSDTENTHYV 556

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2951	100.0	556	AAV35178	Chlamydia pneumoni
2	2951	100.0	556	AA18820	A 60 kDa cysteine-
3	2129.5	72.2	553	AAV37788	Chlamydia trachoma
4	428	14.5	109	AAK07399	Chlamydia Crp gene
5	130.5	4.4	1404	21 AAV59600	Drosophila Seriate
6	130.5	4.4	1483	22 ABG16336	Novel human diagno
7	129.5	4.4	757	22 ABB70052	Drosophila melanog
8	129.5	4.4	1404	14 AAR38304	Sequence of a serr
9	128.5	4.4	1404	22 ABB61998	Drosophila melanog
10	124.5	4.2	1048	22 AAU07143	Chicken CRIM1 prot
11	122	4.1	2931	22 ABB68229	Drosophila melanog

12	120.5	4.1	716	21	AAV76005	Rat Slit homologue
13	120.5	4.1	716	22	AAB55944	Skin cell protein,
14	119.5	4.0	2146	22	ABB62317	Drosophila melanog
15	119	4.0	744	22	ABB58426	Drosophila melanog
16	118.5	4.0	1529	21	AAV76117	Rat Slit homologue
17	118.5	4.0	1529	22	AAB56056	Skin cell protein,
18	117.5	4.0	1503	20	AAV27142	Human slit-2 matur
19	117.5	4.0	1503	20	AAW96701	Slit-like protein
20	117.5	4.0	1525	20	AAV17499	Human slit-1 prote
21	117.5	4.0	1529	20	AAV27145	Human slit-2 prote
22	117.5	4.0	1529	20	AAW96702	Full length slit-1
23	117.5	4.0	1529	20	AAB61239	Human Slit-1 prote
24	115.5	3.9	1036	22	AAU07141	Human CRIM1 protei
25	115.5	3.9	1037	22	AAU07142	Mouse CRIM1 protei
26	115.5	3.9	1810	17	AAU94563	Chicken cytotactin
27	115.5	3.9	2312	22	ABG06402	Novel human diagno
28	114.5	3.9	732	22	AAB61140	Human NOV10 protei
29	114.5	3.9	1036	21	AAV82776	Human chordin rela
30	114.5	3.9	1036	21	AAU53034	Human secreted pro
31	114.5	3.9	1036	22	AAU12242	Human PRO4330 poly
32	114.5	3.9	3478	22	ABG04526	Novel human diagno
33	112.5	3.8	1358	17	AAW99358	Human restrictin.
34	112.5	3.8	1358	18	AAW35743	Recombinant human
35	112.5	3.8	1358	18	AAW18824	B. sphaericus SLP.
36	111.5	3.8	1252	16	AAR80530	Human tenascin-C.
37	111	3.8	2201	22	AAW36935	Slit protein. Xen
38	110.5	3.7	1530	21	AAW28151	Drosophila melanog
39	110	3.7	2703	22	ABG02666	Human protein sequ
40	109	3.7	413	22	AAW92604	Human cytotactin.
41	108.5	3.7	2199	17	AAW94562	Human ADAMTS-1 pro
42	108	3.7	727	20	AAW78435	Human shear stress
43	108	3.7	2705	22	AAW90772	Drosophila melanog
44	108	3.7	3680	22	ABW70878	Human protein SEQ
45	107.5	3.6	1358	22	AAM78544	

ALIGNMENTS

RESULT 1

ID AAV35178 standard; Protein; 556 AA.
XX AAV35178;
AC AAV35178;
XX
DT 13-SEP-1999 (first entry)
XX
Chlamydia pneumoniae transmembrane protein sequence.
XX
De Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW vaccine; neutralising epitope.
XX
Chlamydia pneumoniae.
OS
XX
PN WC9927105-A2.
XX
PD 03-JUN-1999.
XX
PF 20-NOV-1998; 98WO-IB01890.
XX
PR 04-NOV-1998; 98US-0107078.
PR 21-NOV-1997; 97FR-0014673.
(GEST) GENSET.
XX
Griffais R;
XX
WPI; 1999-357842/30.
XX
Genome sequence of Chlamydia pneumoniae
PS Page 1037-1038; Disclosure; 1912pp; English.

```

XX AAY34584-Y35879 represent the proteins encoded by all the open reading
CC frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.
CC C. pneumoniae causes respiratory disease such as pneumonia and
CC bronchitis and is thought to be a contributing factor in heart
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC nodosum or pharyngitis. the polypeptides encoded by the open reading
CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC nucleotides sequences can also be used as immunogenic compositions,
CC especially where the vector directs the expression of a neutralising
CC epitope of C. pneumoniae.
XX
SQ Sequence 556 AA;

Query Match 100.0%; Score 2951; DB 20; Length 556;
Best Local Similarity 100.0%; Pred. No. 4.7e-233;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKLIRRVTVVLTSMASCFASGGTEAAVAESLITKIVASAEKTPAPVPMATAKVRVLR 60
Db 1 msklirrvvtvltasmascfsgsgteaaavaeslittkivasaetkppapvmtakvrivr 60

QY 61 RNKQPVQKSRGAFCDKEFYPCSEGRQCPVEAQOESCYGRLYSVKVNDDCNVEICOSVPE 120
Db 61 rnkpveqksrgafcdkefypceegrcqpveaqqescygrlysvkvnddcnveicqsvpe 120

QY 121 YATVGSPPYPIELTAIGKDKCDVDVVTIQTLPCEAEFVSDDPETPTSDGKLWKIDRLGAG 180
Db 121 yatvgsppypieiltaigkdkcdvdvvtiqtlpceaeefvsddpetptsdgklwkidrlgag 180

QY 181 DKCKITVWVKPLKEGCCFTAAATVCACPELRSYTKCGQPAICIKQEGPDCACLRCPVCYKI 240
Db 181 dkckitvwvkplkegccftaatvcacpelrsytkcgqpaicikqegpdcacrlrcpvcyki 240

QY 241 EVVNTGSAIARNVTVDNPPDGYSHASQORVLSFNLDMPDKKVFTEFCPQRRGOIT 300
Db 241 evvntgsaiaarnvtvdnppdgyshasqorvlsfnldmrdpdkkvftfcpcprrggit 300

QY 301 NVATVTCGCGHKCSANVTTVVNEPCQVNIAGADWSYVCKPVEYSISVSNPGLVLDHVV 360
Db 301 nvatvtcgcghkcsanvttvvnepcqvniagadwsyvcckpveysisvsnpgdlvlhdvv 360

QY 361 IQDTLPSSVTVLEAPGGEICCNKVVRIKEMCPGETIQFLVKAQVPGPFTNOVAVTSE 420
Db 361 iqdtlpssvttvleapggeiccnkvvrrikemcpgetiqflvkaqvgpfpftnoqvavtse 420

QY 421 SNGCTCTSCAETTHWKGLAATHMCLVLDNDPICVGENTVYRICVTNNGSAEDTNVSLIL 480
Db 421 sngctctscactthwkglaaahmclvldndpicvgentvyricvtnngsaedtnvslil 480

QY 481 KFSKELOPIASSGPTKGTISGNTVVFDAIPKLGSKESVEFSVTLKGIAPDARGEAILSS 540
Db 481 kfskelopiasgptkgtisgntvvfdalpklgskesvefsvtlkgiapdargeailss 540

QY 541 DTLTSPVSDTENTHYV 556
Db 541 dltlspvsdtenthyv 556

RESULT 2
AAB18820
ID AAB18820 standard; Protein; 556 AA.
XX
AC AAB18820;
XX
DT 22-JAN-2001 (first entry)
XX
DE A 60 kda cysteine-rich membrane protein of Chlamydia pneumoniae.
XX
KW Cysteine-rich membrane protein; Chlamydia infection; bronchitis;
KW community acquired pneumonia; upper respiratory tract infection; vaccine;

```

```

KW sinusitis.
XX Chlamydia pneumoniae.
XX WO200055326-A1.
XX 21-SEP-2000.
XX
XX 09-MAR-2000; 2000WO-CA00240.
XX 12-MAR-1999; 99US-0123966.
XX (AVET ) AVENTIS PASTEUR LTD.
XX
XX Murdin AD, Omen RP, Wang J, Dunn P;
XX WPI; 2000-618918/59.
XX N-PSDB; AAA75901.
XX
XX New polynucleotides encoding a 60kda cysteine-rich membrane protein
XX from Chlamydia, useful as a vaccine for preventing and treating
XX Chlamydia infection in mammals -
XX
XX Claim 16; Fig 1; 77pp; English.
XX
XX The present sequence represents a Chlamydia 60 kda cysteine-rich membrane
XX protein. The membrane-rich polynucleotide and polypeptide are useful
XX for preventing or treating Chlamydia infection, such as community
XX pneumonia, upper respiratory tract infections, bronchitis and sinusitis.
XX They are also useful for diagnosing Chlamydia infection by assaying a
XX body fluid of a mammal. The polypeptide is useful for vaccine
XX production.
XX
SQ Sequence 556 AA;

Query Match 100.0%; Score 2951; DB 21; Length 556;
Best Local Similarity 100.0%; Pred. No. 4.7e-233;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKLIRRVTVVLTSMASCFASGGTEAAVAESLITKIVASAEKTPAPVPMATAKVRVLR 60
Db 1 msklirrvvtvltasmascfsgsgteaaavaeslittkivasaetkppapvmtakvrivr 60

QY 61 RNKQPVQKSRGAFCDKEFYPCSEGRQCPVEAQOESCYGRLYSVKVNDDCNVEICOSVPE 120
Db 61 rnkpveqksrgafcdkefypceegrcqpveaqqescygrlysvkvnddcnveicqsvpe 120

QY 121 YATVGSPPYPIELTAIGKDKCDVDVVTIQTLPCEAEFVSDDPETPTSDGKLWKIDRLGAG 180
Db 121 yatvgsppypieiltaigkdkcdvdvvtiqtlpceaeefvsddpetptsdgklwkidrlgag 180

QY 181 DKCKITVWVKPLKEGCCFTAAATVCACPELRSYTKCGQPAICIKQEGPDCACLRCPVCYKI 240
Db 181 dkckitvwvkplkegccftaatvcacpelrsytkcgqpaicikqegpdcacrlrcpvcyki 240

QY 241 EVVNTGSAIARNVTVDNPPDGYSHASQORVLSFNLDMPDKKVFTEFCPQRRGOIT 300
Db 241 evvntgsaiaarnvtvdnppdgyshasqorvlsfnldmrdpdkkvftfcpcprrggit 300

QY 301 NVATVTCGCGHKCSANVTTVVNEPCQVNIAGADWSYVCKPVEYSISVSNPGLVLDHVV 360
Db 301 nvatvtcgcghkcsanvttvvnepcqvniagadwsyvcckpveysisvsnpgdlvlhdvv 360

QY 361 IQDTLPSSVTVLEAPGGEICCNKVVRIKEMCPGETIQFLVKAQVPGPFTNOVAVTSE 420
Db 361 iqdtlpssvttvleapggeiccnkvvrrikemcpgetiqflvkaqvgpfpftnoqvavtse 420

QY 421 SNGCTCTSCAETTHWKGLAATHMCLVLDNDPICVGENTVYRICVTNNGSAEDTNVSLIL 480
Db 421 sngctctscactthwkglaaahmclvldndpicvgentvyricvtnngsaedtnvslil 480

QY 481 KFSKELOPIASSGPTKGTISGNTVVFDAIPKLGSKESVEFSVTLKGIAPDARGEAILSS 540

```



```

|||||
Db 115 vpkatvgsppvpeitatkrdcvliitqqlpceaeefvrsdpatttadgklvwkldrl 174
QY 178 GAGDKCKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGOPALCIKQEGPDCACLRCPVC 237
Db 175 gggekstkwvwpkpkkegcftaatvcacpeirsvtkcqqpaicvkegpenaclrcpvc 234
QY 238 YKIEVVNTGSAIARNVTVDNVPDGYSHASGQORVLGFNLDMPKDKKVFTEFCPQRRG 297
Db 235 ykinvvnqgtatarnvvnvvpdsyahssqrvltflgdmqpggehrtitvefcplkr 294
QY 298 QITNATVTCYGGHKCSANVTTVVNEPCVQVNTSGADWSYCKPVEYSISVSNPGLVLH 357
Db 295 ratniamvsycgghkntasvttvinepcvqvsiaagadsvyckpvevsvsnpgdlvlr 354
QY 358 DVVIQDTLPSGVTVLEAPGGEICCNKVVRIKEMCPGETLQFKLVWKAQVPGRETNQAV 417
Db 355 dvvkvdtlspgvtvleagaqiscnkvvwtvkelnpgeslqykvlvraqtpgqftnnvv 414
QY 418 TSENCGCTCTSCAETTTTHWGLAATHMCVLDNDPFCVGENTVYRICVTRNGSAEDTNVS 477
Db 415 kscadcgctscacaeatykwgvaathmcvvdtdcpvcvgentvyrvcvtnrgsaedtnvs 474
QY 478 LILKFSKELQPIASSGPTKGTISGNTVVDALPKLGSKESVEFSVTLKGAIAPGDARGEAI 537
Db 475 lmlkfskelqpsvsgptkgtitgntvvdslprlgsketvefsvtlkavsagdgargai 534
QY 538 LSSDTLTPSVSDTENTHYV 556
Db 535 lssdtltpvsdtenthy 553

RESULT 4
AAR07399 standard; protein; 109 AA.
XX AAR07399;
XX AAR07399;
DT 12-MAR-1991 (first entry)
DE Chlamydia CrP gene prod.
KW PCR; DIANA; solid phase medical diagnosis.
XX Chlamydia trachomatis.
XX WO9011369-A.
PD 04-OCT-1990.
XX 15-MAR-1990; 90WO-EP00454.
XX 22-MAR-1989; 89GB-0006642.
PR 22-MAR-1989; 89GB-0006641.
XX (CEMU-) CEMU BIOTEKNIK.
XX Uhlen M;
XX WPI; 1990-320269/42.
DR N-PSDB; AAQ06256.
XX
PT Amplification of DNA by PCR - using distal DNA on a primer as a
PT handle for attachment to a solid support or a label, used in
PT diagnosis of medical conditions
XX
XX Disclosure; fig 19; 66pp; English.
XX
XX This is the product encoded by the CrP gene. The corresp. gene can
XX be detected utilising DIANA, an improved PCR method with 2 primer
XX pairs.
XX Sequence 109 AA;
SQ

481 kfskelqlassgpkgtisgntvvdalpklgskesvsvtlkgiapgdargeailss 540
541 DTLTPSVSDTENTHYV 556
541 dltltpsvsdtenthy 556

AAY37788
ID AAY37788 standard; Protein; 553 AA.
XX AAY37788;
XX AAY37788;
DT 07-OCT-1999 (first entry)
DE Chlamydia trachomatis cellular envelope protein.
XX Chlamydia trachomatis
XX Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW paratrachoma; inclusion conjunctivitis; genital disease; perihhepatitis;
KW nongonococcal urethritis; epidymitis; cervicitis; salpingitis;
KW Bartholinitis; pneumopathy; venereal lymphogranulomatosis.
XX Chlamydia trachomatis.
XX WO928475-A2.
PD 10-JUN-1999.
XX 27-NOV-1998; 98WO-IB01939.
PR 04-NOV-1998; 98US-0107077.
PR 28-NOV-1997; 97FR-0015041.
PR 17-DEC-1997; 97FR-0016034.
XX (GEST ) GENSET.
XX Griffals R;
XX WPI; 1999-371125/31.
XX
XX Genome sequence of Chlamydia trachomatis
XX Disclosure; Page 1371-1372; 1755pp; English.
XX
XX AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
XX of Chlamydia trachomatis (see AA201425). The polypeptides can be used as
XX vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
XX can also be used to control growth of the microorganism. Chlamydia
XX trachomatis is responsible for a large number of diseases, e.g. eye
XX diseases such as conventional trachoma, nonendemic trachoma,
XX paratrachoma, and inclusion conjunctivitis; genital diseases such as
XX nongonococcal urethritis, epidymitis, cervicitis, salpingitis,
XX perihhepatitis, Bartholinitis; pneumopathy in breast feeding infants;
XX and venereal lymphogranulomatosis. The polypeptides of the invention
XX may be of use in treating these diseases.
XX
XX Sequence 553 AA;
SQ

Query Match 72.2%; Score 2129.5; DB 20; Length 553;
Best Local Similarity 70.7%; Pred. No. 8.8e-166;
Matches 395; Conservative 77; Mismatches 72; Indels 15; Caps 4;

QY 1 MSKLIRRVTVVLTSMASCFASGIEAAVAESLITKIVASAETKPAFPVPMATKRVRLVR 60
Db 7 mnlirravtifavtsvasifasgvltsmaefistnvlsiadtkak--dntshkskar 64
QY 61 RN---KQPEOKSRGAFCDKEFYPCIEGRCQFPVBAQOQESYGRLYSVKVNDDCNVEICQS 117
Db 65 khsketpvnr-----kkvapyeska--tqpkqdcfcgrmytvkvkvdnrdnrvaitqa 114
QY 118 VPEYATVGSPPYPIEILAIAGKKKDCVDVVTQQLPCEAEFVSSDPETPTSDGKLWVKIDRL 177
```


PF 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
PA Dmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
PI N-PSDB; AAS80523.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
PT
XX Claim 20; SEQ ID NO 46695; 103pp; English.
PS
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1483 AA;

Query Match 4.4%; Score 130.5; DB 22; Length 1483;
Best Local Similarity 19.3%; Pred. NO. 0.18;
Matches 118; Conservative 90; Mismatches 217; Indels 185; Gaps 29;

QY 35 ITKIVASAEKTPAPV-----PMTAKKVR-----LVRRNQPVQKSRGAFCDKE--- 78
DB 626 vsqvgsetlhprdlrnfrireqpafrrlnlntllinnq-ikripsgafedlenlk 684
QY 79 FYPCBEGRCQPVAAQOESCYRLYSVKVND--CNVEI---CQSVPEYATVGS----- 126
DB 685 yilykneiqsdrqafglaeqrlrldsnthcdceillwladilktyaesgnaqaai 744
QY 127 -FYPTEILAIGKDCVDDVITQOLCEAEFVSDPPTTSDGKLVKIDRLGAGCKKI 185
DB 745 ceyprri-----qgrsvatitpeelncerpitsepqdaavtsqntvyftcraegnpkpei 800
QY 186 TWVW--KPLKEGCGCTAATVACAPELRSYTKGQPAICIKQEGPCACALRCP--VCYKIEV 242
DB 801 -lwlrnkslerhaaadtgsllsravl-----tcipnkg---slllcptqccgpyk 846
QY 243 VNTGSAIARNVTVDNVPVPGVSHASQRLVSLFNLGDMRPGDKKVFVTEFCPQRRGQITNV 302
DB 847 aktalaaavrvtegsptadhy-----mrpgsephslsltc---rlalpmi 898
QY 303 -----ATVT-----YGGHKCSANVTTVVNEPCVQVNIAGDSWYVC---KPVEYSISVS 349
DB 889 qshlqatltcwlavfltgisvasnvtcll-----ftcfclkgtkvklqvg 933
QY 350 NPGDL-----VLHD--VVIODTLPSGVTVLEA-----PGGEICCNKVVVRIKEMCPGET 396

DB 934 nelsmktdsrlnliddgtlmiqntqetdqqiyqcmaknvagevktqev-----t 982
QY 397 LOFKLVVKAQVPGREF--NOVAVTSESNGCTCTSCAETTHWKGLAATHMVCVLDNDP--- 452
DB 983 lry-----fgspgsflgrvctssvsvctatslqpqntevlgesvtlecatghpppri 1037
QY 453 -----ICVGENTVYRICVTNTRGSAEDTNVSLILKFSK 484
DB 1038 swtrgdrtpdpvdrvnitpsgglyiqnvvgdsgeyacsatnnidsvhatafiiv---q 1094
QY 485 ELQPIASSGPTKGTISGNTVVFD-----ALPKLGSKSESVFSTLKIAPGDAR 533
DB 1095 alpqtvtvpqdrvvlegtdvfqceakgnpppviawtkggslsvd-----r 1141
QY 534 GEATLSSDTL 543
DB 1142 rhlvissgtl 1151

RESULT 7
ABB70052
ID ABB70052 standard; Protein; 757 AA.
XX
AC ABB70052;
XX
DT 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 36948.
DE Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
DR N-PSDB; ABL14155.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 36948; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABBS7737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 757 AA;

Query Match 4.4%; Score 129.5; DB 22; Length 757;

Tue May 28 08:55:16 2002

DE Drosophila melanogaster polypeptide SEQ ID NO 12786.
XX Drosophila: developmental biology; cell signalling, insecticide;
4W pharmaceutical.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
DR N-PSDB; ABL06101.
DR New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX Disclosure: SEQ ID NO 12786; 2lpp + Sequence Listing: English.
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABU16176-ABU130511), expressed DNA
CC sequences (ABU01840-ABU16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1404 AA;
SQ
Query Match 4.4%; Score 128.5; DB 22; Length 1404;
Best Local Similarity 20.5%; Pred. No. 0.24;
Matches 124; Conservative 57; Mismatches 208; Indels 217; Gaps 32;
QY 14 LISMASFASGGTEAAVAESLITKIVASAEETKPAVPMTAKKRVLRNKPQVEQKSRGA 73
Db 543 lttataglnslstallaaltsavastslalqpc-inakecr-----nqp-----gs 590
QY 74 FCDKEFYPCBEGRCQPVEAQOESCYGRLYSVKVNDCNVCILQSVPEYATVGSYPYPIEL 133
Db 591 fa----cicqg-----wggvtcaenlddc-vdgcrn----- 617
QY 134 AIGKKCDVDVITQOLPCEAEFVSSDPE-----TPTSDG-----KLVWK-----IDRLG- 178
Db 618 ---gacldivndyracagsgfkgdcdetdidecatpcnnggcvdmvgkfncicplgy 674
QY 179 AGDKC---KITVMVKPLKEGSCCFTA--ATVCACPELRSYTKCGQ-PAIC----- 221
Db 675 sgsiceakentpspclegchclntpegyychpdragkhceqlrplcsqppcncgcfa 734
QY 222 -----IKQEGPDC-----ACLRCPV---CYKIEVWNTGSAIA 250
Db 735 nvslatsattttttttrkrmakpsglpcsghgscmsdvgtfcc-kehvghtgtfcee 793
QY 251 RNVTVDNVDP-----DG-----YSHASGQRVLS-----FNLGDMRPG-- 282
Db 794 hnnecspncpncggicldgdgdfccmcsgrwtgkrseratgcyagcgnggtcmpgap 853
QY 283 DKKV-----FTVEFCPQRRGQITNVATVTYCGGHKCSANVTTVNPEPCVQVNISGA 333

Db 854 dkalqphrcapgwgtglfcae-----aidqcrqgqchnggt-----ce-----sga 894
QY 334 DW-SYVCKPEVYSISVSNPGDLVLHDVWIQDTLPSTVTVLEAPGGEICCNKVVMRIKEMC 392
Db 895 gwfrvcv-----agfgsgpdcrlnvneccspqcggtatcidgiggygc-----ic 939
QY 393 PGETLOFLVVKVKAQVGRFTNQAVTSESCNGTCTSCAET-----TTHWKGLAATH 443
Db 940 p-----pyrhgicrceillspksacqnasntispytalnrsgnwidialtg 985
QY 444 MCVLDTNDPICVGVNTYRICVTNRGSAEDTNVSLILKFSKELOPIASSGPTKTIGSNT 503
Db 986 rteddenncnaccven-----gtsrctnlwcpncykvdplskssnlsgvckqkhe 1035
QY 504 VVFDAL 509
Db 1036 vcvpael 1041
RESULT 10
AAU07143
ID AAU07143 standard; Protein; 1048 AA.
XX AC AAU07143;
XX AC AC
XX DT 24-OCT-2001 (first entry)
XX DE Chicken CRIM1 protein.
XX CRIM-1; Chicken; human chromosome 2p21-16.3; ophthalmological;
KW neuroprotective; renal; osteopathic; dental; vulnary; immunogen;
KW antibody; gene therapy; neurodegenerative disease; eye disorder;
KW cataract; bone morphogenic protein; BMG; renal disease; bone abnormality;
KW tooth abnormality; wound; S52.
XX OS Gallus gallus.
XX FH Key Location/Qualifiers
FT Region 212..219 /note= "conserved N-terminal motif"
FT Region 348..402 /label= CR_1
FT FT /note= "Cysteine rich repeat"
FT Region 415..468 /label= CR_2
FT FT /note= "Cysteine rich repeat"
FT Region 620..674 /label= CR_3
FT FT /note= "Cysteine rich repeat"
FT Region 691..746 /label= CR_4
FT FT /note= "Cysteine rich repeat"
FT Region 765..820 /label= CR_5
FT FT /note= "Cysteine rich repeat"
FT Region 831..885 /label= CR_6
FT FT /note= "Cysteine rich repeat"
XX WO200138519-A1.
XX 31-MAY-2001.
XX 24-NOV-2000; 2000WO-AU01435.
XX 26-NOV-1999; 99AU-0004348.
XX (UYQU) UNIV QUEENSLAND.
XX Little M, Yamada T, Holmes G, Georgas K, Kolle G, Wilkinson L;
PI WPI; 2001-343951/36.
DR

Tue May 28 08:55:16 2002

AAV76119 are proteins with an N-terminal signal sequence, indicating that they are secreted. Sequences AAY75986-Y75989, AAY76061-Y76071, AAY76106-Y76109 and AAY76121-Y76122 are proteins with one or more putative transmembrane domains.

Sequence 716 AA;

Query Match 4.1%; Score 120.5; DB 21; Length 716;
Best Local Similarity 20.4%; Pred. No. 0.42; Indels 191; Gaps 24;
Matches 92; Conservative 42; Mismatches 126;

72 GAFCDKEFPCEGRGQ-----PYEAQOQESQ-----YGRLY-----SVKVNDD 109
313 gehcdidfdcdqknckngahctdavnngytcvcpgysglfcefppmvflrtspcdnfd 372
110 C-----NVEICOSPEYATVGSPPYPIEILAIKDKDC-----VDVVITQQLPCEAE 154
373 cnggaqcilrvnepicqlpgy-----lgek-ceklsvsliivnkesylqip 418
155 FVSSDPET-----TPTSDGKLWKIDRLGAGDKCKITVMVKPLKEGCGCTTAATVCACP 207
419 sakvrpqtntltqiatdedsqillyk-----gdkdhav-----eslegi----- 458
208 ELRGYTKGQPAICIKQEGPDCACLRCPVCYKIEVWTGS-----AIAKNVTVDNP 258
459 -rasydtgshpasal-----ysvetindgnfhiwelltdsslsldvsg 502
259 VPDGYSHASGQVLSFN-----LGDMPGDKKVFVEFCPQRGRGQITNVATVYTCGGHKCS 314
503 spkiitniskstlnfdsplyvggm-pgknnvaslrqapqng-----tsfhqci 551
315 ANVTWVNEPCVQVNIAGDMSYVCKPVEYSISNPGDLVLDVHVVIODTLPSGVTVLEA 374
552 rh-----lyinseiqdfrkv--pmqtgi-----l 573
375 PGGEICCNKVVWRKEMCPGETLQFKLVVKAQVPGRTNQAVTSESCNGTCTCAETTT 434
574 pgcepchkkv-----cangtgcq-----pssqsgficece----- 603
435 HWKGLAATHMCVLDNDPICVGVNTVYRICV 465
604 gwmq-----plcdqrdndp-clgnkcvhgtcl 629

RESULT 13
AAB55944
ID AAB55944 standard; Protein: 716 AA.
AC AAB55944;
XX 08-MAR-2001 (first entry)
DT Skin cell protein, SEQ ID NO: 183.
DE Rat; skin cell; cytostatic; antinflammatory; anti-HIV;
KW neutropic; neuroprotective; vulnery; immunomodulatory; vaccine;
KW keratinocyte growth stimulation; cancer; angiogenesis inhibition;
KW inflammation; neurological disease.
OS Rattus sp.
XX WO2000069884-A2.
XX 23-NOV-2000.
XX 15-MAY-2000; 2000WO-NZ00075.
XX 14-MAY-1999; 99US-0312283.
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX Watson JD, Strachan L, Onrust R, Sleeman M, Kumble KD, Murison JG;
PI

1473 csqgkneysitpdpntclg-----clstddprcatnqtltnicsqipyte 1521
368 GVTVEAPGGEI--C-----CNKVVWRKEMCPGETLQFKLVYKAQVPGRTNQ 414
1522 cvtdqidsqntirgcmsslsgddfyecltgdgkicetctgercngisvfpad--rrkcyg 1579
415 VAVTSESCNGTCTCAETTHWKLAAATHMCVLDNDPICV-----IASSGPTK 496
1580 cnstsdnpcaaspatlest-----vcpiysqdescvttllnglthrgcscslltcs 1629
456 --GENTVYRICVTRNGSAEDTNVSLILKFSKELQP-----IASSGPTK 496
1630 dpsdartervc-----ssadgcntinlekigedfgpawdvpieclvcsdtdcassggtl 1685
497 GTRISGN-----TWFD--ALPKLGSKEV 518
1686 tkcsydnvcvtvfdsgsvtqrqesv 1713

RESULT 12
AAY76005
ID AAY76005 standard; Protein: 716 AA.
XX
AC AAY76005;
XX 27-MAR-2000 (first entry)
DT Rat Slit homologue, SEQ ID NO:183.
DE Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
KW secreted; transmembrane; inflammation; cancer; neurological disease;
KW angiogenesis; tumour vascularisation; growth disorder;
KW developmental disorder; skin wound; hair follicle disorder;
KW anti-inflammatory; cytostatic; neuroprotective; vulnery.
XX Rattus sp.
XX WO9955865-A1.
XX 04-NOV-1999.
XX 29-APR-1999; 99WO-NZ00051.
XX 29-APR-1998; 98US-0069726.
XX 09-NOV-1998; 98US-0188930.
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;
PI
XX WPI: 2000-072177/06.
XX N-PSDB: AA261669.
XX Novel polynucleotides useful for the treatment of various conditions
XX including wounds and cancer -
XX Claim 4; Page 125-126; 235pp; English.

The invention relates to novel nucleic acid sequences derived from rat dermal papilla, human keratinocytes and neonatal foreskin fibroblasts, and mouse embryonic skin, keratinocyte stem cells and transit amplifying cells. Polypeptides of the invention may be used to treat inflammation, cancer and neurological diseases. The proteins may be used to stimulate the growth and motility of keratinocytes, to inhibit the growth of cancer cells, to modulate angiogenesis and tumour vascularisation, to modulate skin inflammation, to modulate epithelial cell growth and to inhibit binding of HIV-1 to leukocytes. The invention may also be used to treat growth and developmental defects, skin wounds and hair follicle disorders. Sequences AAY75942-Y76123 represent polypeptides encoded by cDNA sequences derived from several mouse, rat or human skin cell types. Sequences AAY75942-Y75947, AAY76020-Y76021, AAY76094-Y76104 and

```
XX WPI: 2001-007495/01.
DR N-PSDB; AAC99602.
XX
XX New isolated polynucleotide used in the identification of genetic
PT disorders and encoding polypeptides used for treating inflammatory
PT disease, cancer and neurological diseases -
XX
XX Claim 4; Page 161-163: 352pp; English.
XX
XX The present sequence is a polypeptide which is expressed in
CC mammalian skin cells. The polypeptide is useful for stimulating
CC keratinocyte growth and motility, inhibiting the growth of cancer cells,
CC modulating angiogenesis, inhibiting angiogenesis and vascularisation of
CC tumours, modulating skin inflammation, stimulating the growth of
CC epithelial cells, inhibiting the binding of human immunodeficiency virus
CC (HIV)-1 to leukocytes, and treating inflammatory disease, cancer and
CC neurological diseases. The polynucleotide can be used as a marker, in
CC the identification of genetic disorders, and for the design of
CC oligonucleotides for examining expression patterns.
XX
XX Sequence 716 AA;
SQ
Query Match 4.18; Score 120.5; DB 22; Length 716;
Best Local Similarity 20.4%; Pred. No. 0.42;
Matches 92; Conservative 42; Mismatches 126; Indels 191; Gaps 24;
QY 72 GAFCDKEFYPCEE-----PVEAQOESC---YGRLY-----SVKVNDD 109
Db 313 gehcdldfdcdnkcngahctdavnqytcvcepgysglfcesppmvflrtspcdnfd 372
QY 110 C-----NVEICQSVPEYATVGSYPPIEILAIAGKKDC-----VDVVITQQLPCEAE 154
Db 373 cqnagacilrvnepicqlqpsy-----lgeek-cekivsvliivnesylqip 418
QY 155 FVSDPET-----TPTSDGKLWVKIDRLGAGDKCKITVWVKPLKGGCCFTAAATVCACP 207
Db 419 sakvrptnitiqatdedsgllyk-----gdkdhiav-----eslegi----- 458
QY 208 ELRSYTKCGQPAICIKQEGPCACLRCPVCYKIEVWNTGS-----AIARNVTVDNP 258
Db 459 -rasydtgshpsai-----ysvetindnfnhivelitldssislsvdgg 502
QY 259 VPDGYSHASQVLSFN-----LGMRRPGDKKVFVEFCPPORRGQITNVATVTCGGHKCS 314
Db 503 spkiitnlskqstlnfdsplyvggm-pgknnavslrqapqng-----tsfhgci 551
QY 315 ANVTTVVNEPCVOVNISGADMSYVCKPVEYSISVSNPGDLVLHDVVIQDTLPISGVTVLEA 374
Db 552 rn-----lyinselqdfkrv--pmqtgi-----l 573
QY 375 PGGEICCNKVMRIKEMCPGETLQPKLVYKAQVFCRETNQAVTSENCGTCTSCAETTT 434
Db 574 pgcepchkv-----cahgctq-----pssqsqfteece--- 603
QY 435 HWKGLAATHMVCVLDNDPTVCENTVYRICV 465
Db 604 gwmg---plcdqrtdnp-clgnkcvhgtcl 629
RESULT 14
ABB62317
ID ABB62317 standard; Protein; 2146 AA.
XX
XX ABB62317;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 13743.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
```

```
XX Drosophila melanogaster.
OS WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI: 2001-656860/75.
XX
XX N-PSDB; ABL06420.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Disclosure; SEQ ID NO 13743; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 2146 AA;
SQ
Query Match 4.0%; Score 119.5; DB 22; Length 2146;
Best Local Similarity 20.3%; Pred. No. 2.3;
Matches 74; Conservative 33; Mismatches 121; Indels 137; Gaps 19;
QY 72 GAFCDKEFYPCEE-----GRC-----QPVEAQOESCY--GRLY 102
Db 419 gafqtnvdecnknpelnggrcfdtygywylcqcldgwgeicdrpmtcdtqqclnggtcl 478
QY 103 SVKVNDDC-----NVEICQSVPEYATVGSYPPIEILAIAGKKDCVDVVIQDTLPCEAEFV 156
Db 479 dkpifgqlcpepytgelcqlapsca---qqcpldsecvvgk-cvckpsssgyncc--- 530
QY 157 SSDPETPTSDGKLWVKIDRLGAGDKCKITVWVKPLKGGCCFTAAATVCAC----- 206
Db 531 -----tstgdgasalaltpl-----ncnatng-kcinggtcsmngtheycavvgysgdrce 579
QY 207 -PELRSYTKCGQPAICIKQEGPCACLRCPVCYKIEVWNTGSAIARNVTVDNPDPGYSH 265
Db 580 kaencsplncqepmncvqng---clcpnkvc-----nqcatcpqng--- 619
QY 266 ASGQRVLSFNGLDMRPGD-----KKVTFVEFCPPORRGQITN 301
Db 620 --gecv-----dlpngdyeckctrgwtgctcndvdectlhpkicngiclknekygk 671
QY 302 VATVTYCGGHKCSANVTTVVNEPCVOVNISGAD-----WSYVCKP-----VEYSI 346
Db 672 yctpgftgvh-cdsdvdecisfpc-----Ingatchnklnayecvcqpgyegencevidde 726
QY 347 SVSNP 351
Db 727 cgsnp 731
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

QM protein - protein search, using sw model

Run on: May 25, 2002, 20:42:06 : Search time 54.34 Seconds
(without alignments)
249.920 Million cell updates/sec

Title: US-09-523-647-2
Perfect score: 2951
Sequence: 1 MSKLIRRVTVLALTMASC.....ILSSDTLTSPVSDPTENTHVV 556

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2425594 residues
Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	435	14.7	109	1	US-08-477-270-20	Sequence 20, Appl
2	141	4.8	25	4	US-09-025-596-38	Sequence 38, Appl
3	130.5	4.4	1404	2	US-08-400-159-2	Sequence 2, Appl
4	130.5	4.4	1404	3	US-08-611-729A-2	Sequence 2, Appl
5	120.5	4.1	771	4	US-09-188-930-183	Sequence 183, App
6	119	4.0	20	4	US-09-025-596-41	Sequence 41, Appl
7	117.5	4.0	1525	3	US-09-191-647-2	Sequence 2, Appl
8	117.5	4.0	1525	4	US-09-540-245A-2	Sequence 2, Appl
9	117.5	4.0	1525	4	US-09-540-153-2	Sequence 2, Appl
10	115.5	3.9	1810	5	PCT-US95-11684-4	Sequence 4, Appl
11	112.5	3.8	1358	1	US-08-404-665-4	Sequence 4, Appl
12	112.5	3.8	1358	1	US-08-404-671-4	Sequence 4, Appl
13	112.5	3.8	1222	2	US-08-404-781-4	Sequence 15, Appl
14	111.5	3.8	1222	2	US-08-682-517-15	Sequence 9, Appl
15	111.5	3.8	1252	2	US-08-682-517-9	Sequence 9, Appl
16	106.5	3.7	2199	5	PCT-US95-11684-2	Sequence 84, Appl
17	106.5	3.6	4654	4	US-08-476-515A-84	Sequence 84, Appl
18	106.5	3.6	4655	4	US-08-652-877-84	Sequence 86, Appl
19	106.5	3.6	4655	4	US-08-652-877-86	Sequence 88, Appl
20	106.5	3.6	4655	4	US-08-652-877-88	Sequence 90, Appl
21	106.5	3.6	4655	4	US-08-652-877-90	Sequence 90, Appl
22	106	3.6	19	4	US-09-025-596-39	Sequence 39, Appl
23	105.5	3.6	2813	3	US-08-896-449A-2	Sequence 2, Appl
24	105.5	3.6	2813	3	US-09-132-652-2	Sequence 2, Appl
25	105	3.6	2556	1	US-08-083-590A-20	Sequence 20, Appl
26	105	3.6	2556	1	US-08-532-384-20	Sequence 20, Appl
27	104	3.5	886	3	US-09-110-116-3	Sequence 3, Appl

28	104	3.5	2703	1	US-08-185-432-19	Sequence 19, Appl
29	103.5	3.5	2471	1	US-08-185-432-16	Sequence 16, Appl
30	103.5	3.5	2471	1	US-08-083-590A-19	Sequence 19, Appl
31	103.5	3.5	2471	3	US-08-532-384-19	Sequence 19, Appl
32	102	3.5	2523	1	US-08-185-432-18	Sequence 18, Appl
33	101.5	3.4	1480	3	US-09-191-647-7	Sequence 7, Appl
34	101.5	3.4	1480	4	US-09-540-245A-7	Sequence 7, Appl
35	101.5	3.4	1480	4	US-09-540-153-7	Sequence 7, Appl
36	101.5	3.4	1480	5	PCT-US91-09055-2	Sequence 2, Appl
37	100	3.4	2556	1	US-08-185-432-17	Sequence 17, Appl
38	99	3.4	1940	2	US-08-644-271-30	Sequence 30, Appl
39	98.5	3.3	2647	2	US-08-583-562B-8	Sequence 8, Appl
40	98.5	3.3	2647	2	US-08-779-113-8	Sequence 10, Appl
41	97.5	3.3	1193	3	US-08-400-159-10	Sequence 10, Appl
42	97.5	3.3	1193	3	US-08-611-729A-10	Sequence 5, Appl
43	97.5	3.3	1480	4	US-09-182-024A-5	Sequence 2, Appl
44	97.5	3.3	2050	2	US-08-347-594A-2	Sequence 2, Appl
45	96.5	3.3	1833	3	US-08-479-722B-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-477-270-20
: Sequence 20, Application US/08477270
: Patent No. 5629158
: GENERAL INFORMATION:
: APPLICANT: UHLEN, Mathias
: TITLE OF INVENTION: SOLID PHASE DIAGNOSIS OF MEDICAL
: TITLE OF INVENTION: CONDITIONS
: NUMBER OF SEQUENCES: 30
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 1800 Diagonal Road, Suite 500
: CITY: Alexandria
: STATE: Virginia
: COUNTRY: USA
: ZIP: 22313-0299
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/477,270
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: PRIOR APPLICATION NUMBER: US/08/261,010
: FILING DATE:
: APPLICATION NUMBER: US 07/781,157
: FILING DATE: 07-NOV-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 16787/153 DFBC
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703)836-9300
: TELEFAX: (703)683-4109
: TELEX: 899149
: INFORMATION FOR SEQ ID NO: 20:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 109 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-477-270-20

Query Match 14.7%; Score 435; DB 1; Length 109;
Best Local Similarity 72.5%; Pred. No. 6.7e-33;
Matches 79; Conservative 16; Mismatches 14; Indels 0; Gaps 0;

RESULT 3
 US-08-400-159-2
 : Sequence 2, Application US/08400159
 : Patent No. 5869282
 : GENERAL INFORMATION:
 : APPLICANT: Ish-Horowicz, David
 : APPLICANT: Henrique, Domingos M.P.
 : APPLICANT: Lewis, Julian H.
 : APPLICANT: Myat, Anna M.
 : APPLICANT: Fleming, Robert J.
 : APPLICANT: Artavanis-Tsakonas, Spyridon
 : APPLICANT: Mann, Robert S.
 : APPLICANT: Gray, Grace E.
 : TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
 : TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
 : NUMBER OF SEQUENCES: 20
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Pennie & Edmonds
 : STREET: 1155 Avenue of the Americas
 : CITY: New York
 : STATE: New York
 : COUNTRY: USA
 : ZIP: 10036-2711
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:

RESULT 4
US-08-611-729A-2
; Sequence 2, Application US/08611729A
; Patent No. 6004924

Tue May 28 08:55:17 2002

GENERAL INFORMATION:
 APPLICANT: Ish-Horowitz, David
 APPLICANT: Henrique, Domingos M.P.
 APPLICANT: Lewis, Julian H.
 APPLICANT: Myat, Anna M.
 APPLICANT: Fleming, Robert J.
 APPLICANT: Artavanis-Tsakonas, Spyridon
 APPLICANT: Mann, Robert S.
 APPLICANT: Gray, Grace E.
 TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
 TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/611,729A
 FILING DATE: 06-MAR-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Misrock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 7326-037
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-9741/8864
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1404 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-611-729A-2

Query Match 4.4%; Score 130.5; DB 3; Length 1404;
 Best Local Similarity 20.5%; Pred. No. 0.0041;
 Matches 124; Conservative 57; Mismatches 208; Indels 217; Gaps 32;
 QY 14 LTSMASCFASGGIEAAVAESLITKIVASAEKTPAPVMTAKKVLRRNKPVEQKSRGA 73
 DB 543 LTTTAAITGSLSTALLAALTSVAVSTSLAIGPC-INAKCR-----NQP-----GS 590
 QY 74 FCDKEYPCCEGRCPVEAQESCYGRYSVKVNDNCNVICQSPVEATVGSPPYIEL 133
 DB 591 FA-----CICKEG-----WGGVTCAEMLDDC-VGQCRN----- 617
 QY 134 AIGKDCVDVITQQLPCAEFVSSDPE-----TTPTSDG-----KLVWK-----IDRLG- 178
 DB 618 ---GATCIDLVNDYRCACASGFTGRDCETDIDECATSPCRNGECVDMVGKENCICPLGY 674
 QY 179 AGDKC---KITVMVKPLKECCCFTA--ATVCACPELRSYTKCQ-PAIC----- 221
 DB 675 SGLUCEEAKENCTPSCLEGHCLNTPEGYCHCPDPRAGKHCBQLRPLCSQPPCNGGCTA 734
 QY 222 -----IKOEGPDC-----ACLRCPV---CYKIEVVNTGSAIA 250
 DB 735 NVSLATSATTTTTTTTATTTRKMAKPSGLPCSHGSCMSDVGTFC-KCHVGHGTGTCF 793
 QY 251 RNVTVDNVPP-----DG-----YSHASQORVLS-----ENLGDMPRG-- 282
 DB 794 HNLNCSNPNCRNGICLDGGDFTCBMSGWTGKRCERATGCGYAGCQCGCTCMPGAP 853

QY 283 DKKV-----FTVEFCPORRGQITNVAITVTCYGGHKCSANVTTVVNEPCVQVNISGA 333
 DB 854 DKALQPHCRCAPGWTGLFCAE-----AIDQCRGQPCNGGT-----CE-----SGA 894
 QY 334 DW-SYVCKEVEYSISNPGDLVLDHVVYIQTLPSPGVTVLEAPGGEICCNKVVRHKEWC 392
 DB 895 GWFRCVC-----AQFSGPDCRINYNECSPOPCQGATCIDGIGGYS-IC 939
 QY 393 PGETLQFKLVVKAQVPGRETNOVAVTSESNGCTCTSCAET-----TTHWKGLAATH 443
 DB 940 P-----PGRHGLRCEILLSDPKSACQNASNTISPYTALNRSONWLDIALTG 985
 QY 444 MCVLDNDPCVGVNTVIRICVTNRGSAEDTNVSLTLKFKSELOPIASSGPTKGTISGNT 503
 DB 986 RTEDDENCNACVEN-----GTSRCTNLWCGLPNCYKVDPLSKSNLSGVCKQHE 1035
 QY 504 VVFDAL 509
 DB 1036 VCVPAL 1041
 RESULT 5
 US-09-188-930-183
 ; Sequence 183, Application US/09188930A
 ; Patent No. 6150502
 ; GENERAL INFORMATION:
 ; APPLICANT: Watson, James D.
 ; APPLICANT: Strachan, Lorna
 ; APPLICANT: Sleeman, Matthew
 ; APPLICANT: Onrust, Rene
 ; APPLICANT: Murlison, James Greg
 ; TITLE OF INVENTION: Compositions Isolated From Skin Cells
 ; TITLE OF INVENTION: and Methods For Their Use
 ; FILE REFERENCE: 11000.1011c1
 ; CURRENT APPLICATION NUMBER: US/09/188,930A
 ; CURRENT FILING DATE: 1998-11-09
 ; NUMBER OF SEQ ID NOS: 348
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 183
 ; LENGTH: 771
 ; TYPE: PRT
 ; ORGANISM: Rat
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: (717)...(717)
 ; US-09-188-930-183

Query Match 4.1%; Score 120.5; DB 4; Length 771;
 Best Local Similarity 20.4%; Pred. No. 0.014;
 Matches 92; Conservative 42; Mismatches 126; Indels 191; Gaps 24;
 QY 72 GAFCDKEYPCCEGRCP-----PVEAQOESC---YGRLY-----SVKYNDD 109
 DB 313 GEHCIDIDFDDQDNKCKNGAHCCTDAVNGYTCVCPGEGYSGLFCEFPMPVFLRTSPCNFD 372
 QY 110 C-----NVFICQSPVEATVGSPPYIEILAIAGCKDC-----VDVITQQLPCEAE 154
 DB 373 CONGAQCIIRVNEPICQCLPGY-----LGEK-CEKLVSVSILVKNESYLQIP 418
 QY 155 FVSSDPET-----TPTSDGKLVWKIDRLGAGDKCKITVWPKLKECCCFATATVCACP 207
 DB 419 SAKVRPQTNITLQIATDEDSGLLYK-----GDKDHIAV--BSIEGI----- 458
 QY 208 ELRSYTKCQQAICIKQEGPCACLRCPVCKYIEVVNTGS-----ATARNVTVDNP 258
 DB 459 -RASVDTGSHPASAI-----YSVETINDGNFHVILLTLDSSLISLSDVG 502
 QY 259 VPDGISHASQORVLSFN-----LGDMPRGKVKVTFVEFCFQRRGQITNVAITVTCYGGHKCS 314
 DB 503 SPKIIITNLKSQSTLNFDSPLYVGGM-PGKNNVASLRQAFCQNG-----TSFHGCI 551
 QY 315 ANVTTVVNEPCVQVNISGADWSYVCKPVEYSISVSNPGLVLDHVVYIQTLPSPGVTVLEA 374

Db 552 RN-----LYINSELQDFRKV--PMQTGI-----L 573
Qy 375 PGEICCNKVVRIKEMCPGETLQFLKLVVKAQVGRFTNQAVTSESNCGTCTSCAETTT 434
Db 574 PCEPCHKKV-----CAHGTGQ-----PSSQSGFTCECEE--- 603
Qy 435 HWKGLAATHMVCVLDNDPICVGTNVRICV 465
Db 604 GWMG-----PLCDQRTNDP-CLGNKCVHGTCL 629

RESULT 6

US-09-025-596-41
; Sequence 41, Application US/09025596
; Patent No. 6340463
; GENERAL INFORMATION:
; APPLICANT: Mitchell, William M.
; APPLICANT: Stratton, Charles W.
; TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE
; FILE REFERENCE: VDB98-01
; CURRENT APPLICATION NUMBER: US/09/025,596
; EARLIER FILING DATE: 1998-02-18
; EARLIER FILING DATE: 1997-08-14
; EARLIER FILING DATE: 1997-08-14
; EARLIER FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 41
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-025-596-41

Query Match 4.0%; Score 119; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.5e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 418 TSESNGCTCSAETTHWK 437
Db 1 TSESNGCTCSAETTHWK 20

RESULT 7

US-09-191-647-2
; Sequence 2, Application US/09191647
; Patent No. 6046015
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey
; APPLICANT: Kid, Thomas
; APPLICANT: Brose, Katja
; TITLE OF INVENTION: Modulating Robo: Ligand Interactions
; FILE REFERENCE: B98-031-3
; CURRENT APPLICATION NUMBER: US/09/191,647
; EARLIER FILING DATE: 1998-11-13
; EARLIER FILING DATE: 1997-11-14
; EARLIER FILING DATE: 1997-11-14
; EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1525
; TYPE: PRT
; ORGANISM: human
US-09-191-647-2

Query Match 4.0%; Score 117.5; DB 3; Length 1525;

Best Local Similarity 20.3%; Pred. No. 0.074;
Matches 92; Conservative 40; Mismatches 125; Indels 197; Gaps 23;
Qy 72 GAFCDKEFYCEGRGQ-----PVEAQQESC---YGRLY-----SYKVND 109
Db 1066 GEHCDIDFDDCQDNKCKNGAHCDAVNGYTCICPEGYSGLCFCEFPMPVLPRTSPCDNFD 1125
Qy 110 C-----NVEICQSVPEYATVGSPIEILATGKDKCDVW-----ITQQLP 150
Db 1126 CONGAQCIIVRINEPICQCLPGYQGEKLVSVNFINKESYLQIPSAKVRPQTNITLQI- 1184
Qy 151 CEAEFVSSDPETPTSDGKLWNKIDRLGAGDKCKIIVWVKPLKEGCCFTAAATVCACPEL- 209
Db 1185 -----ATDEDSGILLYK-----GDKDHIAV-----ELY 1207
Qy 210 -----RSYTKGGOAICIKOEGPDCACLRCPVCYKIEVNTGS-----AI--ARNVTV 255
Db 1208 RGRVRASYDTGSHPSAI-----YSVETINDGNFHVIVELLALDQSLSLV 1252
Qy 256 DNPVDPGYSHASQORVLSFN-----LGDMPRGKKVFTVEFCPQRRGQITNVAITYCGH 311
Db 1253 DGGNPKIITNLKQSTLNFDSPLYVGM-PGKSNVASLRAQPGQNG-----TSFH 1301
Qy 312 KCSANVTTVNPECVQVNIAGADWSYVCKPVEYSISVNPGLVLJHDVVIQDTLPSTV 371
Db 1302 GCIRN-----LYINSELQDFRKV--PMQTGI----- 1325
Qy 372 LEAPGGEICCNKVVRIKEMCPGETLQFLKLVVKAQVGRFTNQAVTSESNCGTCTSCAE 431
Db 1326 --LFGCEPCHKKV-----CAHGTGQ-----PSSQAGFTCEQOE 1356
Qy 432 TTHWKGAAATHMVCVLDNDPICVGTNVRICV 465
Db 1357 --GWMG-----PLCDQRTNDP-CLGNKCVHGTCL 1382

RESULT 8

US-09-540-245A-2
; Sequence 2, Application US/09540245A
; Patent No. 6270984
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey
; APPLICANT: Kid, Thomas
; APPLICANT: Brose, Katja
; APPLICANT: Tessier-Lavigne, Marc
; TITLE OF INVENTION: Modulating Robo: Ligand Interactions
; FILE REFERENCE: B98-031-3
; CURRENT APPLICATION NUMBER: US/09/540,245A
; CURRENT FILING DATE: 2000-03-31
; PRIOR FILING DATE: 1997-11-14
; PRIOR FILING DATE: 1997-11-14
; PRIOR FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1525
; TYPE: PRT
; ORGANISM: human
US-09-540-245A-2

Query Match 4.0%; Score 117.5; DB 4; Length 1525;
Best Local Similarity 20.3%; Pred. No. 0.074;
Matches 92; Conservative 40; Mismatches 125; Indels 197; Gaps 23;

Qy 72 GAFCDKEFYCEGRGQ-----PVEAQQESC---YGRLY-----SYKVND 109
Db 1066 GEHCDIDFDDCQDNKCKNGAHCDAVNGYTCICPEGYSGLCFCEFPMPVLPRTSPCDNFD 1125
Qy 110 C-----NVEICQSVPEYATVGSPIEILATGKDKCDVW-----ITQQLP 150
Db 1126 CONGAQCIIVRINEPICQCLPGYQGEKLVSVNFINKESYLQIPSAKVRPQTNITLQI- 1184

Tue May 28 08:55:17 2002

Db 1253 DGNPKIITNLSKOSTLNFDSPLYVGGM-PGKSNVASLRQAPGONG-TSPH 1301
QY 312 KCSANVTTVVNEPCVOVNIISGADWSYVCKPVEYSISVSNPGDLVLDVVDVQDLPSCVTV 371
Db 1302 GCIRN-----LYINSELQDFQKV--PMQTGI-----1325
QY 372 LEAPGEICCNKVVWRKEMCPGETLQFKLVVKAQVPGREFTNOVAVTSESCGTCTSCAE 431
Db 1326 --LPGCEPCHKV-----CAHGTCQ-----PSSOAGFTCECQE 1356
QY 432 TTHWKGLAATHCMCVLDTNDPICVGVNTVYRICV 465
Db 1357 ---GWMG-----PLCDQRTNDP-CLGNKCVHGTCL 1382
RESULT 10
PCT-US95-11684-4
; Sequence 4, Application PC/TUS9511684
; GENERAL INFORMATION:
; APPLICANT: THE SCRIPPS RESEARCH INSTITUTE
; TITLE OF INVENTION: CYTOTACTIN DERIVATIVES THAT STIMULATE
; TITLE OF INVENTION: ATTACHMENT AND NEURITE OUTGROWTH, AND METHODS OF MAKING
; NUMBER OF SEQUENCES: 28
; NUMBER OF SEQUENCES: AND USING SAME
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; STREET: 10666 North Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: PCT/US95/11684
; FILING DATE: 14-SEP-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/308,359
; FILING DATE: 16-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Logan, April C. 33,950
; REGISTRATION NUMBER: BEC0019P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1810 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-11684-4

QY 151 CEAEFVSSDPETPTSDGKLVKIDRLGAGDKCKITVWVKPLKRGCCFTAAATVCACPEL- 209
Db 1185 -----ATDEDSGILLYK-----GKDHIAV-----ELY 1207
QY 210 -----RSYTKGQPAICIKQEGPDCACILRCPVCYKIEVVNTGS-----AI--ARNVTV 255
Db 1208 RGRVRASYDTGSHPASAI-----YSVETINDGNFHVIELLALDQSLSLV 1252
QY 256 DNPVDPGVSHASGORVLSFN-----LQDMRPGDKKVTVEFCPORRGQTNNVATVYCCGGH 311
Db 1253 DGNPKIITNLSKOSTLNFDSPLYVGGM-PGKSNVASLRQAPGONG-----TSPH 1301
QY 312 KCSANVTTVVNEPCVOVNIISGADWSYVCKPVEYSISVSNPGDLVLDVVDVQDLPSCVTV 371
Db 1302 GCIRN-----LYINSELQDFQKV--PMQTGI-----1325
QY 372 LEAPGEICCNKVVWRKEMCPGETLQFKLVVKAQVPGREFTNOVAVTSESCGTCTSCAE 431
Db 1326 --LPGCEPCHKV-----CAHGTCQ-----PSSOAGFTCECQE 1356
QY 432 TTHWKGLAATHCMCVLDTNDPICVGVNTVYRICV 465
Db 1357 ---GWMG-----PLCDQRTNDP-CLGNKCVHGTCL 1382
RESULT 9
US-09-540-153-2
; Sequence 2, Application US/09540153
; Patent No. 6270995
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey
; APPLICANT: Kid, Thomas
; APPLICANT: Brose, Katja
; TITLE OF INVENTION: Modulating Robo: Ligand Interactions
; FILE REFERENCE: B98-031-3
; CURRENT APPLICATION NUMBER: US/09/540,153
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/191,647
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 60/081,057
; PRIOR FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 2
; LENGTH: 1525
; TYPE: PRT
; ORGANISM: human
US-09-540-153-2
Query Match 4.0%; Score 117.5; DB 4; Length 1525;
Best Local Similarity 20.3%; Pred. No. 0.074; Indels 197; Gaps 23;
Matches 92; Conservative 40; Mismatches 125; Indels 197; Gaps 23;
QY 72 GAFCDREFYCEGRQ-----PVEAQOESC-----YGRLY-----SVKVNDD 109
Db 1066 GEHCIDIDFDDCKDNKNGAHCHTDAVNGYTCICPEGYSLGFCFSPMVLPRTPSCDNFD 1125
QY 110 C-----NVEICOSVPEYATVGPYPPIELAIKDKCDVDV-----ITQQLP 150
Db 1126 CONGAOCIVRINEPICQLPGYQCEKELVSVNFINKESYLOIPSAKVRPNTNLIQI- 1184
QY 151 CEAEFVSSDPETPTSDGKLVKIDRLGAGDKCKITVWVKPLKRGCCFTAAATVCACPEL- 209
Db 1185 -----ATDEDSGILLYK-----GKDHIAV-----ELY 1207
QY 210 -----RSYTKGQPAICIKQEGPDCACILRCPVCYKIEVVNTGS-----AI--ARNVTV 255
Db 1208 RGRVRASYDTGSHPASAI-----YSVETINDGNFHVIELLALDQSLSLV 1252
QY 256 DNPVDPGVSHASGORVLSFN-----LQDMRPGDKKVTVEFCPORRGQTNNVATVYCCGGH 311

QY 183 CKITVWVKPLKGGCCGTAATVC-----ACPELRSYTKCQPAICIKQE-----GP 227
Db 444 CO-----QGRGICNGQCECHGFTGDCGELRCPNDCNHCRCVNGQCVCDGEGTGE 495
QY 228 DCACLRCP-VYKIEVNTGSAIARNVTVDNVPDGY-----SHASGORVLS 273
Db 496 DCGELRCPNDCH-----NRCRCVBCRCVCDN-----GFMGDCGELSCPNDCQHGRVCV-- 544
QY 274 FNLGDMRPGDKKVTVEFCPQRGQIITNATVTCYCGHKCSANVTTVVNEPCV-QVNISG 332
Db 545 -----DGRVCVCHGFTGEDCRERS-----C-PNDCN-NVGRVCVEGRCVEEYMG 587
QY 333 ADNSYCKPVEYSISVSNPGLDVL-----HDVVI-----QDTLP---S 367
Db 588 IDCSDVSPPT--GLTNTVNTDKTNLEWKHENLVNEVLYTVYVTSGLGLDLQFTVPGNQ 645
QY 368 GVTVLE-APGGE-----ICCNK-----VWRIKEMCPG-ETIQFLVVKVAQVPR-- 410
Db 646 SATIHELEGVYFIRFALLKNNKIPVSARVATYLPAPGLKFKRSVRETSVQVEWDPL 705
QY 411 -----FTNOVAVTSESNGCTCTSCAETTHWKGLA-----ATHMCVLDNDPI 453
Db 706 SISFDGWLNVFRNMOKDDNGDITSLKRPETSYMQPGLAPGOQYVNSLHIVKNNTRGP- 764
QY 454 CVGENTVY--RICVTNRGSAED-TNVSLLKFSKELOPIASSGPTKG--TISGNTVYVD- 507
Db 765 --GLSRVITTKLDAPSQIEAKDVTDTALITWSKPLAETIEGIELTYGPKDVPDGRITDL 822
QY 508 -----ALPKLGSKESVEFSVTL---KGIAPGDARGEAILSSDTLTSP-----VSDTEN 552
Db 823 SEDENOYSIGNLRPHTEYEYVTLISRRGDMESDPAKEVFYTD--LDAPRNLKRVSDQN 880
QY 553 T 553
Db 881 S 881

RESULT 11
US-08-404-665-4
; Sequence 4, Application US/08404665
; Patent No. 5591583
; GENERAL INFORMATION:
; APPLICANT: Reid, Robert A.
; APPLICANT: Ackley, Rhonda L.
; APPLICANT: Hemperly, John J.
; TITLE OF INVENTION: HUMAN RESTRICTIN AND NUCLEIC ACID
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
; STREET: 1 Becton Drive
; CITY: Franklin Lakes
; STATE: NJ
; COUNTRY: US
; ZIP: 07417
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/404,665
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fugit, Donna R.
; REGISTRATION NUMBER: 32,135
; REFERENCE/DOCKET NUMBER: P-3341
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1358 amino acids
; TYPE: amino acid

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-404-665-4

Query Match 3.88; Score 112.5; DB 1; Length 1358;
Best Local Similarity 19.28; Pred. No. 0.18;
Matches 125; Conservative 60; Mismatches 233; Indels 233; Gaps 28;

QY 23 SGGTEAAVAPSLITKIIVASAEITKPAVPMTAKKVLVRRNKQPVQKSRGAFCDKEFPYPC 82
Db 80 SSGLE-----ASAEQE-----VSAEDETALAEYMGQTSQVTHRINFPK 122
QY 83 EGRGCPVEAQOESCGRLYSVK-----VNDNCNVEICQSVPEYATVGYSPYIILAIGK 137
Db 123 KACPCSSAQLQELLSEMLELREVERSVLRDQCNAACQ-----ESAATGQ 168
QY 138 KDCVDVVIITQOLPCEAEFVSSDPETPTSDGKLVMKIDRLGA-----GDKCKITVWVK 190
Db 169 LDVI-----PHCSGHGNFSEFCGCGICNCGWFGKNCSEPYC-- 204
QY 191 PL---KEGCGCTAATVC-----ACPELRSYTKCQPAICIKQE-----GPDCACL 232
Db 205 PLOCCSRGVGVGQCCICDSEYSGDDCSELRCPTDCSSRGLCVGDCVCEEPYTGDCREL 264
QY 233 RCP-VCYKIEVYNTGSAIARNVTVDNVPDGYSHASGORVLSFNLDGMRPGDKKVTVEF 291
Db 265 RCPGDCSGKRCATGTCLCE-----EGY-----VGED 291
QY 292 CPQR-----RGQITNVATVYCG--GHKCSANVTTVVNEPCVOVNISG-----AD 334
Db 292 CGORCLNACSGRGQCEGLCVCEGYQGPCSAVA-----PPEDLRVAGISDRSIELE 345
QY 335 WSYCKPVEYISVSNPGDLVLDVVIODTLP---SGVTLEAPG---EICCCKVWVRI 388
Db 346 WDGPMATVEYVSYQ---PTALGGLOLQORVPCDWSGVTTITTELEPLTYNISVYAVISNI 402
QY 389 -----KEMCPGETLQKLV-----VKAQV 407
Db 403 LSLPITAKVATHLSTPOGLQKFTITETTVQWEPFSEFQDGEISFIPKNEGVIQAV 462
QY 408 PGRFT-----NOVAVTSESNGCTCTSCAETTHWKGLAATHMCVLDNDP 452
Db 463 PSDVTSFNQTLGKPEEYIVNVVALKEQARSPTTSASVSTVID-----GPTOILVRDVS 518
QY 453 ICVGENTVYRICVTNRGSAEDTNVSLILKF-----SKELOPIASSGPTKGITSGN 502
Db 519 VAFVE-----WIPPRAKVD-----FILLKYGLVGGEGRTTFRLOPPLSQVSVQALRPGS 568
QY 503 TVVFDALPKLGSKESVEFSVTLKGIAPGDARGEAILSSDTLTSPVSDTEN 553
Db 569 RYEVSVSAVRGTNES--DSATTQTTTIDAPKLNLRVGSRTATSLDLEWDNS 617

RESULT 12
US-08-404-671-4
; Sequence 4, Application US/08404671
; Patent No. 5635360
; GENERAL INFORMATION:
; APPLICANT: Reid, Robert A.
; APPLICANT: Ackley, Rhonda L.
; APPLICANT: Hemperly, John J.
; TITLE OF INVENTION: HUMAN RESTRICTIN AND NUCLEIC ACID
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
; STREET: 1 Becton Drive
; CITY: Franklin Lakes
; STATE: NJ
; COUNTRY: US

ZIP: 07417
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/404.671
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fugit, Donna R.
REGISTRATION NUMBER: 32,135
REFERENCE/DOCKET NUMBER: P-3341
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1358 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-404-671-4

Query Match 3.8%; Score 112.5; DB 1; Length 1358;
Best Local Similarity 19.2%; Pred. No. 0.18;
Matches 125; Conservative 60; Mismatches 233; Indels 233; Gaps 28;

Qy 23 SGGIEAAVAESLITKIVASAEKPAVPMTAKKVLRRNKQPVQKSRGAFCDKEFYPC 82
Db 80 SSGLE-----ASAEQE-----VSAEDELAEYMGQTSDEHSQVTFTHRNFPK 122
Qy 83 EGRGCPVEAQOESCGRLYSVK-----VNDNCNVEICQSVPEYATVGSYPPIELAIK 137
Db 123 KACPSSSAQVQLLESLRIEMLEREVSVLRQCNCANCCQ-----ESATGQ 168
Qy 138 KCDVDVVTIQLPCEAEFVSSDPETPTSDGKLWKIDRLGA-----GDKCKITVWVK 190
Db 169 LDYI-----PHCSGHGNEFSESCGICNEGWFKNKSEPYC-- 204
Qy 191 PL---REGCFTAAATVC-----ACPELRSYTKCGQPAICIKOE-----GPDCACL 232
Db 205 PLGCSRGVCDGQCICDSEYSGDDCSELRCPTDCSSRGLCVGECVCEPYTGDCREL 264
Qy 233 RCP-VCYKIEVNTGSAIARNVTVDNVPDGYSHASGORVLSFNLGDMRPGDKKVFVEF 291
Db 265 RCPGDSGKRCATGTCICE-----EGY-----VGED 291
Qy 292 CPQR-----RGQITNATVTCG---GHKCSANVTTVNNEPCVQVNISSG-----AD 334
Db 292 CGORQCLNACSGRGQCEGLCVCEGYGQPCSAVA-----PPEDLRVAGISDRSIELE 345
Qy 335 WSYVCKPVEYSISVSNPGLVLDHVVIOQTLP---SGVTVLEAPGG---EICCNKVVWRI 388
Db 346 WDGPMAVTEYVISYQ---PTALGGQLQORVPGDWSGVITTELEPGLTYNISVYAVISNI 402
Qy 389 -----KEMCPGETLOFLV-----VKAQV 407
Db 403 LSLPITAKVATHLSTPGLOFRTITETTVVEVQWEPFSEFSDGWEISIPKNNEGGVIAQV 462
Qy 408 PGRFT-----NOVAVTSESNGCTCTSCAETHHWKGLAATHMCVLDTNDP 452
Db 463 PSDVTSFNTGLKPGBEYIVNVVAKLEQARSPTTSASVSTVID---GPTQILVRVSDT 518
Qy 453 ICVENTVYRICVTNRGSAEDTNVSLIKF-----SKELOPIASSSGPTKGTISGN 502
Db 519 VAFV-----WIPPRAKVD---FILLKYLGVGEGGRTTFRLOPLSQYSVQALRPGS 568
Qy 503 TVVFDALPKLGSKESEVFSVTKGIAPDARGEAFLSSDITLTSVDSVTENT 553
Db 569 RYEVSVAVRGTNES--DSAITQFTTEIDAPKNLRVGSRTATSLDLEWMS 617

RESULT 13
US-08-404-781-4
Sequence 4, Application US/08404781
Patent No. 5681931
GENERAL INFORMATION:
APPLICANT: Reid, Robert A.
APPLICANT: Ackley, Rhonda L.
APPLICANT: Hemperly, John J.
TITLE OF INVENTION: HUMAN RESTRICTIN AND NUCLEIC ACID
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
STREET: 1 Becton Drive
CITY: Franklin Lakes
STATE: NJ
COUNTRY: US
ZIP: 07417
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/404.781
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Fugit, Donna R.
REGISTRATION NUMBER: 32,135
REFERENCE/DOCKET NUMBER: P-3341
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1358 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-404-781-4

Query Match 3.8%; Score 112.5; DB 1; Length 1358;
Best Local Similarity 19.2%; Pred. No. 0.18;
Matches 125; Conservative 60; Mismatches 233; Indels 233; Gaps 28;

Qy 23 SGGIEAAVAESLITKIVASAEKPAVPMTAKKVLRRNKQPVQKSRGAFCDKEFYPC 82
Db 80 SSGLE-----ASAEQE-----VSAEDELAEYMGQTSDEHSQVTFTHRNFPK 122
Qy 83 EGRGCPVEAQOESCGRLYSVK-----VNDNCNVEICQSVPEYATVGSYPPIELAIK 137
Db 123 KACPSSSAQVQLLESLRIEMLEREVSVLRQCNCANCCQ-----ESATGQ 168
Qy 138 KCDVDVVTIQLPCEAEFVSSDPETPTSDGKLWKIDRLGA-----GDKCKITVWVK 190
Db 169 LDYI-----PHCSGHGNEFSESCGICNEGWFKNKSEPYC-- 204
Qy 191 PL---REGCFTAAATVC-----ACPELRSYTKCGQPAICIKOE-----GPDCACL 232
Db 205 PLGCSRGVCDGQCICDSEYSGDDCSELRCPTDCSSRGLCVGECVCEPYTGDCREL 264
Qy 233 RCP-VCYKIEVNTGSAIARNVTVDNVPDGYSHASGORVLSFNLGDMRPGDKKVFVEF 291
Db 265 RCPGDSGKRCATGTCICE-----EGY-----VGED 291
Qy 292 CPQR-----RGQITNATVTCG---GHKCSANVTTVNNEPCVQVNISSG-----AD 334
Db 292 CGORQCLNACSGRGQCEGLCVCEGYGQPCSAVA-----PPEDLRVAGISDRSIELE 345
Qy 335 WSYVCKPVEYSISVSNPGLVLDHVVIOQTLP---SGVTVLEAPGG---EICCNKVVWRI 388
Db 346 WDGPMAVTEYVISYQ---PTALGGQLQORVPGDWSGVITTELEPGLTYNISVYAVISNI 402

QY 389 -----KMCPCGETLQFKLV-----VKAQV 407
: : : : :
Db 403 LSLPITAKVATHLSTPQGLQFKTITETTVQVQWEPFSPFDGWEISFPKNNEGGVIAQ 462
: : : : :
QY 408 PKRFT-----NOVAVTSNCGTCTSCAETTHWKGLAATHMCAVLDNDP 452
: : : : :
Db 463 PSDVTSFNGTLKPGEEYIVNVVAKLQCARSPPTSASVIVID---GPQILVRVSDT 518
: : : : :
QY 453 ICVGENTYVRIQVTHRGSAEDTNVSLIKF-----SKELOPIASSGPTKGIISGN 502
: : : : :
Db 519 VAFVE-----WIPRAKVD-----FILLKYLGVGEGGRTRFQPPLSYVQALRPGS 568
: : : : :
QY 503 TVVFEDALPKLGSKEVSFVTLKGIAPGDARGEAILSDTLTSPVSDTENT 553
: : : : :
Db 569 RYEVSVSAVRGTNES--DSATTOFTTEIDAPKNLRVGSRTATSLDLEWDS 617
: : : : :
RESULT 14
US-08-682-517-15
: Sequence 15, Application US/08682517
: Patent No. 5874267
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Expression of surface layer proteins
: NUMBER OF SEQUENCES: 25
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/682,517
: FILING DATE:
: CLASSIFICATION:
: INFORMATION FOR SEQ ID NO: 15:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1222 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-682-517-15
Query Match 3.8%; Score 111.5; DB 2; Length 1222;
Best Local Similarity 20.6%; Pred. No. 0.19;
Matches 114; Conservative 65; Mismatches 209; Indels 165; Gaps 24;
QY 99 GRLYSVKVNDDCNVEICQSVPEYATVGSYPPIEILAIKDKCDVDVITQOLPCEAFVSS 158
||| : : : : :
Db 523 GREYKIVKD-----KAGNLAKNEIVNVAFNEDKDRVIS--TVTNKAFVDT 566
: : : : :
QY 159 DPET-----TPTSD-GKLWVKIDRLGAGDKCKITVWV----KPLKEGCGCFTA 200
: : : : :
Db 567 DPDTAVYFTGDKAKOISVKTNKDGATEVIGSDTVNDYATPIAWIDINTSDAKOGLDEG 626
: : : : :
QY 201 ATVCACP-----ELRSYTKCGOPAIKIQEGPCACLRCPVCYKIEVNV---244
: : : : :
Db 627 EPKAVAPISYFOAPYLDGSAIKAYKKSDLNKAFTKFDGSETA-----VFAAELVNSGK 680
: : : : :
QY 245 --TGSALARNVTVDPVPGYSHASQORVLSNGLDMRPGDKKYFTVEFCPOBGGITNV 302
||| : : : : :
Db 681 KVTGTSIKK-----ATYIYNTGANDIKVDNQVISPNSRYTVTYE 720
: : : : :
QY 303 ATVTYCGGHKCSA---NVTVV-----VNEPCVQVNISGADWSYCKPVEYSISVSNP 351
: : : : :
Db 721 ATLSSTGTVTTPAKNLEVTSDGKTTAVKVIATGIAVNTDGDYAKFAATAFTATFATNE 780
: : : : :
QY 352 GDVLVHDVVIQ-DTLPSSG-----VTVLEAPGGEICCNKVVWR 387
: : : : :
Db 781 VPNSYTGATQNTADSGSNSIWFAGKNPVKYAGVSGTKYKYGANGNEV-FGEAAWE 839
: : : : :
QY 388 -----IKEMCPGETLQFKLVKRAQVPGRTNOVAVTS-----ESNCGTCTSC 429
: : : : :
Db 870 ALLTQYATEGOKVTSISYVNDGDTVTTFKVISAVNSSTEAIKPVAPTTGAULTLTPA 929
: : : : :
US-08-682-517-9
: Sequence 9, Application US/08682517
: Patent No. 5874267
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Expression of surface layer proteins
: NUMBER OF SEQUENCES: 25
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/682,517
: FILING DATE:
: CLASSIFICATION:
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1252 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-682-517-9
Query Match 3.8%; Score 111.5; DB 2; Length 1252;
Best Local Similarity 20.6%; Pred. No. 0.2;
Matches 114; Conservative 65; Mismatches 209; Indels 165; Gaps 24;
QY 99 GRLYSVKVNDDCNVEICQSVPEYATVGSYPPIEILAIKDKCDVDVITQOLPCEAFVSS 158
||| : : : : :
Db 553 GREYKIVKD-----KAGNLAKNEIVNVAFNEDKDRVIS--TVTNKAFVDT 596
: : : : :
QY 159 DPET-----TPTSD-GKLWVKIDRLGAGDKCKITVWV----KPLKEGCGCFTA 200
: : : : :
Db 587 DPDTAVYFTGDKAKOISVKTNKDGATEVIGSDTVNDYATPIAWIDINTSDAKOGLDEG 656
: : : : :
QY 201 ATVCACP-----ELRSYTKCGOPAIKIQEGPCACLRCPVCYKIEVNV---244
: : : : :
Db 657 EPKAVAPISYFOAPYLDGSAIKAYKKSDLNKAFTKFDGSETA-----VFAAELVNSGK 710
: : : : :
QY 245 --TGSALARNVTVDPVPGYSHASQORVLSNGLDMRPGDKKYFTVEFCPOBGGITNV 302
||| : : : : :
Db 711 KVTGTSIKK-----ATYIYNTGANDIKVDNQVISPNSRYTVTYE 750
: : : : :
QY 303 ATVTYCGGHKCSA---NVTVV-----VNEPCVQVNISGADWSYCKPVEYSISVSNP 351
: : : : :
Db 751 ATLSSTGTVTTPAKNLEVTSDGKTTAVKVIATGIAVNTDGDYAKFAATAFTATFATNE 810
: : : : :
QY 352 GDVLVHDVVIQ-DTLPSSG-----VTVLEAPGGEICCNKVVWR 387
: : : : :
Db 811 VPNSYTGATQNTADSGSNSIWFAGKNPVKYAGVSGTKYKYGANGNEV-FGEAAWE 869
: : : : :
QY 388 -----IKEMCPGETLQFKLVKRAQVPGRTNOVAVTS-----ESNCGTCTSC 429
: : : : :
Db 870 ALLTQYATEGOKVTSISYVNDGDTVTTFKVISAVNSSTEAIKPVAPTTGAULTLTPA 929
: : : : :
US-08-682-517-9
: Sequence 9, Application US/08682517
: Patent No. 5874267
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Expression of surface layer proteins
: NUMBER OF SEQUENCES: 25
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/682,517
: FILING DATE:
: CLASSIFICATION:
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1252 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-682-517-9

Db 840 ALLTQYATEGOKVTSISYVNDGDTVTTFKVISAVNSSTEAIKPVAPTTGAULTLTPA 899
: : : : :
QY 430 A-----ETTHHKGLAATHMCAVLDNDPFCVGVNTYVYRICVTHRGSAEDTNVSLIL---480
: : : : :
Db 900 AGGLVLDLTATNTLIGIS---LADADLN---VSATVTVDTATVSLKDSA--NNSLSLTLVET 951
: : : : :
QY 481 ----KESKELOPIASSGPTKGIISGNTVVF--DALPKLGSKEVSFVTLKGIAPGDARG 534
: : : : :
Db 952 GANTGVFATTVQAGTLSLTAGTL---TVTYADAKNAAGVAENITASVTLK-----KT 1001
: : : : :
QY 535 EAILSSDTLSPV 547
: : : : :
Db 1002 TGAITSDTFTQGV 1014
: : : : :
RESULT 15
US-08-682-517-9
: Sequence 9, Application US/08682517
: Patent No. 5874267
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Expression of surface layer proteins
: NUMBER OF SEQUENCES: 25
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/682,517
: FILING DATE:
: CLASSIFICATION:
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1252 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-682-517-9
Query Match 3.8%; Score 111.5; DB 2; Length 1252;
Best Local Similarity 20.6%; Pred. No. 0.2;
Matches 114; Conservative 65; Mismatches 209; Indels 165; Gaps 24;
QY 99 GRLYSVKVNDDCNVEICQSVPEYATVGSYPPIEILAIKDKCDVDVITQOLPCEAFVSS 158
||| : : : : :
Db 553 GREYKIVKD-----KAGNLAKNEIVNVAFNEDKDRVIS--TVTNKAFVDT 596
: : : : :
QY 159 DPET-----TPTSD-GKLWVKIDRLGAGDKCKITVWV----KPLKEGCGCFTA 200
: : : : :
Db 587 DPDTAVYFTGDKAKOISVKTNKDGATEVIGSDTVNDYATPIAWIDINTSDAKOGLDEG 656
: : : : :
QY 201 ATVCACP-----ELRSYTKCGOPAIKIQEGPCACLRCPVCYKIEVNV---244
: : : : :
Db 657 EPKAVAPISYFOAPYLDGSAIKAYKKSDLNKAFTKFDGSETA-----VFAAELVNSGK 710
: : : : :
QY 245 --TGSALARNVTVDPVPGYSHASQORVLSNGLDMRPGDKKYFTVEFCPOBGGITNV 302
||| : : : : :
Db 711 KVTGTSIKK-----ATYIYNTGANDIKVDNQVISPNSRYTVTYE 750
: : : : :
QY 303 ATVTYCGGHKCSA---NVTVV-----VNEPCVQVNISGADWSYCKPVEYSISVSNP 351
: : : : :
Db 751 ATLSSTGTVTTPAKNLEVTSDGKTTAVKVIATGIAVNTDGDYAKFAATAFTATFATNE 810
: : : : :
QY 352 GDVLVHDVVIQ-DTLPSSG-----VTVLEAPGGEICCNKVVWR 387
: : : : :
Db 811 VPNSYTGATQNTADSGSNSIWFAGKNPVKYAGVSGTKYKYGANGNEV-FGEAAWE 869
: : : : :
QY 388 -----IKEMCPGETLQFKLVKRAQVPGRTNOVAVTS-----ESNCGTCTSC 429
: : : : :
Db 870 ALLTQYATEGOKVTSISYVNDGDTVTTFKVISAVNSSTEAIKPVAPTTGAULTLTPA 929
: : : : :
US-08-682-517-9
: Sequence 9, Application US/08682517
: Patent No. 5874267
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Expression of surface layer proteins
: NUMBER OF SEQUENCES: 25
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/682,517
: FILING DATE:
: CLASSIFICATION:
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1252 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-682-517-9

Tue May 28 08:55:19 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

QM protein - protein search, using sw model

Run on: May 25, 2002, 20:48:46 ; Search time 56.2 seconds
(without alignments)
950.635 Million cell updates/sec

Title: US-09-523-647-2
Perfect score: 2951
Sequence: 1 MSKLRRVVVLTALTSMA SCFASGGIEAAVAESLTKIVASAEKTPAPVMTAKKRLVLR 556

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2951	100.0	556	1 S12602	60K cysteine-rich
2	2951	100.0	556	2 A86560	60 kDa Cysteine-rich
3	2553.5	86.5	557	1 B39439	60K cysteine-rich
4	2488	84.3	558	2 JC5204	60K cysteine-rich
5	2176	73.7	554	2 C81671	60 kDa outer membr
6	2162.5	73.3	547	2 B43584	60K cysteine-rich
7	2160.5	73.2	553	2 D71515	60K cysteine-rich
8	2129.5	72.2	547	1 A32244	60K cysteine-rich
9	151.5	5.1	1474	2 F69009	probable membrane surface protein 51
10	135	4.6	2233	2 T28669	hypothetical prote
11	130.5	4.4	756	2 D75622	gene serrate prote
12	130.5	4.4	1408	2 S16148	protein-glutamine conserved hypothet
13	129	4.4	772	2 A48822	hypothetical prote
14	127.5	4.3	998	2 T25878	hypothetical prote
15	127	4.3	693	2 T34264	crumbs protein - f
16	124	4.2	2195	2 A35672	tenascin - eastern
17	123.5	4.2	2139	2 A43902	gip1 protein precu
18	123	4.2	647	2 A32501	hypothetical prote
19	123	4.2	1295	2 A32501	otogelin - mouse
20	122	4.1	491	2 T42214	hypothetical prote
21	122	4.1	2910	2 D71170	hypothetical prote
22	121	4.1	641	2 T22274	Januskin precursor,
23	120.5	4.1	354	2 A45445	alpha-51D immobili
24	120	4.1	1356	2 A45445	alpha-51D-immobili
25	120	4.1	2533	2 T28675	hypothetical prote
26	120	4.1	2533	2 T28674	probable peptidogl
27	119	4.0	1599	2 T16210	restrictin precurs
28	118	4.0	1612	2 AB1347	
29	117.5	4.0	1353	1 JH0675	

30	117.5	4.0	1658	2 D75489	hypothetical prote
31	116.5	3.9	565	2 T16408	hypothetical prote
32	115.5	3.9	577	2 A60501	thrombomodulin pre
33	115.5	3.9	1810	1 A32230	tenascin precursor
34	115	3.9	2543	2 T31687	surface antigen - p
35	114.5	3.9	2918	2 A54105	fibrillin-2 precur
36	113.5	3.8	626	2 T27319	hypothetical prote
37	113.5	3.8	2946	2 T15840	hypothetical prote
38	113	3.8	2907	2 A57278	secreted leucine-r
39	112.5	3.8	1025	2 T42626	hypothetical prote
40	112.5	3.8	1620	2 T27283	hypothetical prote
41	112.5	3.8	1806	2 T23298	hypothetical prote
42	112.5	3.8	2271	2 F90073	hypothetical prote
43	112	3.8	2471	2 A49128	cell-fate determin
44	112	3.8	2871	2 A55624	fibrillin-1 precur
45	111.5	3.8	1746	1 S19694	tenascin precursor

ALIGNMENTS

RESULT 1

S12602

60K cysteine-rich outer membrane protein precursor [similarity] - Chlamydoiphila pneum
C:Species: Chlamydoiphila pneumoniae, Chlamydia pneumoniae
C:Date: 30-Sep-1993 #sequence-revision 27-Jun-1994 #text-change 11-May-2000
C:Accession: S12602; H72063; A81604
R:Watson, M.W.; Al-Mahdawi, S.; Lamden, P.R.; Clarke, I.N.
Nucleic Acids Res. 18, 5299, 1990
A:Title: The nucleotide sequence of the 60kDa cysteine rich outer membrane protein of
A:Reference number: S12602; MUID:90384850
A:Accession: S12602
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-556 <WAT>
A:Cross-references: GB:X53511; NID:g550564; PIDN:CAA37590.1; PID:g550566
A:Experimental source: isolate IOL-207
R:Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606
A:Accession: H72063
A:Molecule type: DNA
A:Residues: 1-556 <ARN>
A:Cross-references: GB:AE001640; GB:AE001363; NID:g4376845; PIDN:AAD18697.1; PID:g437
A:Experimental source: strain CWL029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A:Reference number: A81500; MUID:20150255
A:Accession: A81604
A:Molecule type: DNA
A:Residues: 1-556 <REA>
A:Cross-references: GB:AE002180; GB:AE002161; NID:g7189117; PIDN:AAF38068.1; PID:g718
A:Experimental source: strain AR39, HL cells
C:Comment: This protein is associated with the differentiation of reticulate bodies
. It may also be an important virulence factor.
C:Genetics:
A:Gene: OmcB; CP0195
C:Superfamily: 60K cysteine-rich outer membrane protein
C:Keywords: membrane protein; virulence
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-40/Domain: propeptide #status predicted <PRO>
F:41-556/Product: 60K cysteine-rich outer membrane protein #status predicted <MAT>

Query Match 100.0%; Score 2951; DB 1; Length 556;
Best Local Similarity 100.0%; Pred. No. 2.9e-197;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 MSKLRRVVVLTALTSMA SCFASGGIEAAVAESLTKIVASAEKTPAPVMTAKKRLVLR 60
|||||

Db	121	YATVGSPPDIEILAIKQKCDVVVITQQLPCEAEFVSSDPETPTSDCKLWIKIDRLGAG	180
Qy	181	DKCKITVMVKPLKEGECCTAATVCACPELRSYTKGOPALCIKOEKPGDCACLRCPVCYKI	240
Db	181	DKCKITVMVKPLKEGECCTAATVCACPELRSYTKGOPALCIKOEKPGDCACLRCPVCYKI	240
Qy	241	EVVNTGSAIARNVTVDNPVDPDGYSHASGQRVLISFNIGDMRPGDKKVFTEFCFQRRGOIT	300
Db	241	EVVNTGSAIARNVTVDNPVDPDGYSHASGQRVLISFNIGDMRPGDKKVFTEFCFQRRGOIT	300
Qy	301	NVATVTCGGHKCSANVTTVNPECPQVNIISGADWSYVCKPVEYSISVSNPGLDLVLDVV	360
Db	301	NVATVTCGGHKCSANVTTVNPECPQVNIISGADWSYVCKPVEYSISVSNPGLDLVLDVV	360
Qy	361	IODTLPSGVTVLEAPGGEICCNVWRIKEMCPGETLQFKLVYKAQVGRFTNOVAVTSE	420
Db	361	IODTLPSGVTVLEAPGGEICCNVWRIKEMCPGETLQFKLVYKAQVGRFTNOVAVTSE	420
Qy	421	SNCGTCTSCAETTHWKGLAATHMCLVDNDPICVGENTVYRICVNTNRGSAEDTNVSLIL	480
Db	421	SNCGTCTSCAETTHWKGLAATHMCLVDNDPICVGENTVYRICVNTNRGSAEDTNVSLIL	480
Qy	481	KFSKELOPIASSGPTKGTISGNTVVFDALPKLGSKESEFVSVTLKGIAPGDAKGEAILSS	540
Db	481	KFSKELOPIASSGPTKGTISGNTVVFDALPKLGSKESEFVSVTLKGIAPGDAKGEAILSS	540
Qy	541	DTLTSPPVSDTENTHVVY	556
Db	541	DTLTSPPVSDTENTHVVY	556
RESULT	3		
B39439			
60K cysteine-rich outer membrane protein 1 precursor - Chlamydomophila psittaci			
C:Species:	Chlamydomophila psittaci	Chlamydia psittaci	
C:Date:	21-Feb-1992	#sequence_revision	27-Jun-1994 #text_change 31-Mar-2000
C:Accession:	B39439	S12603	
R:Everett,	K.D.E.; Hatch,	T.P.	
J. Bacteriol.	173,	3821-3830,	1991
A:Title:	Sequence analysis and lipid modification of the cysteine-rich envelope pro		
A:Reference number:	A39439	MUID:91267949	
A:Accession:	B39439		
A:Molecule type:	DNA		
A:Residues:	1-557	<EVE>	
A:Cross-references:	GB:M6116; NID:g144489; PIDN:AAB61619.1; PID:g144491		
A:Experimental source:	strain 6BC		
R:Watson,	M.W.; Lambden,	P.R.; Clarke,	I.N.
Nucleic Acids Res.	18,	5300,	1990
A:Title:	The nucleotide sequence of the 60KDa cysteine rich outer membrane protein c		
A:Reference number:	S12603; MUID:90384851		
A:Accession:	S12603		
A:Status:	not compared with conceptual translation		
A:Molecule type:	DNA		
A:Residues:	1-44, 'A', '46-72, 'E', '74-557	<SWAT>	
A:Cross-references:	GB:X53512; NID:g40625; PIDN:CAA37592.1; PID:g40627		
A:Experimental source:	strain EAE/A22/M		
C:Function:			
A:Description:	associated with the differentiation of reticulate bodies into element		
O	be an important virulence factor.		
C:Superfamily:	60K cysteine-rich outer membrane protein		
C:Keywords:	membrane protein; virulence		
F:1-22/Domain:	signal sequence #status predicted <SIG>		
F:23-40/Domain:	propeptide #status predicted <PRO>		
F:41-557/Product:	60K cysteine-rich outer membrane protein 1 #status predicted <MAT>		
Query Match	86.5%	Score 2553.5;	DB 1; Length 557;
Best Local Similarity	85.3%	Pred. No. 1.1e-169;	
Matches 475;	Conservative	51;	Indels 1; Gaps 1;
QY	1	MSKLIRRVVTVLALTSNASCFSAGGIEAAVAESLITKIVASAEKTPAPV-PMTAKKVRILV	59

Db 1 MSKLRRVVTVLALSMASFFASGKIEAAAAESLATRTASTENSDNDNVFQATAKKVREG 60
QY 60 RRNKQVEQKSRGAFCDKFEFPCCEGRGCPVEAQOESCYGRLYSVKVNDDCNVEICQSP 119
Db 61 RRNKQVEQKHTGAFCDKFEFPCGGGOCQPDVATQESCYGKMYCVRVNDDCNVEISQSP 120
QY 120 EVATVGSPPYPIELAIAGKKDCVDVVTITQOLPCEAEFVSSDPETPTSDGKLWVKIDRLGA 179
Db 121 EVATVGSPPYPIELAVGKDCVNVVITQOLPCEVEFVSSDPATPTSDSKLIWTIDRLGQ 180
QY 180 GDCKKITVWVKPLKEGCCFTAAATVCACPELRSYTKGQPAICIKQEGPCDCAELRCPCVYK 239
Db 181 GECKKITVWVKPLKEGCCFTAAATVCACPELRSYTKGQPAICIKQEGPECAELRCPCVYK 240
QY 240 IEVNTGSAIARNVTVNDNPVPGYSHASQORVLSFNLGDMRPGDKVFTVEFCPORRGOI 299
Db 241 IEVNTGSAIARNVTVNDNPVPGYSHASQORVLSFNLGDMRPGDKVFTVEFCPORRGOI 300
QY 300 TNVATVYCGGHKCSANVTTVNNEPCQVNVISGADWSYVCKPVEYSISVSNPGDLVLHDV 359
Db 301 TNVATVYCGGHKCSANVTTVNNEPCQVNVISGADWSYVCKPVEYTIIVSNPGDLKLYD 360
QY 360 VIQDTLPSTGVTLEAPGGEICCNKVMRIKEMCPGETLQFKLVVKAQVPGFTNOVAVTS 419
Db 361 VIEDTAPSGATILEAAGAEICCNKAVMCIKEMCPGETLQFKLVVKAQVPGFTNOVAVTS 420
QY 420 ESNCGTCTSCAETHHMKGLAATHMCLVDTNDPVCVGENTVYRICVTVNRGSAEDTNVSLI 479
Db 421 NSDCGTCTSCAETHHMKGLAATHMCLVDTNDPVCVGENTVYRICVTVNRGSAEDTNVSLI 480
QY 480 LKFSKELOPIASSGPTKGTISGNTVVFDPALPKLGSKESEFVSITLKGAPDARGEAIL 539
Db 481 LKFSKELOPVSSSGPTKGTITGNTVVFDPALPKLGSKESEFVSITLKGAPDARGEAIL 540
QY 540 SDTLTSPVSDTENTHVV 556
Db 541 SDTLTVPVADTENTHVV 557
RESULT 4
JC5204
60K cysteine-rich outer membrane protein 2 precursor - Chlamydomophila psittaci
C:Species: Chlamydomophila psittaci, Chlamydia psittaci
C:Date: 20-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 31-Mar-2000
C:Accession: JC5204
R:Hisia, R.; Bavoil, P.M.
Gene 176, 155-162, 1996
A:Title: Sequence analysis of the omp2 region of Chlamydia psittaci strain GPIC: Structu
A:Reference number: JC5203; MUID:97075924
A:Accession: JC5204
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-558 <HSI>
A:Cross-references: GB:U41759; NID:g1783376; PIDN:AAB41143.1; PID:g1783382
A:Experimental source: strain GPIC
C:Genetics:
A:Gene: omp2
C:Superfamily: 60K cysteine-rich outer membrane protein
C:Keywords: membrane protein; virulence
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-40/Domain: propeptide #status predicted <PRO>
F:41-558/Product: 60K cysteine-rich outer membrane protein 2 #status predicted <MAT>
Query Match 84.38; Score 2488; DB 2; Length 558;
Best Local Similarity 82.1%; Pred. No. 4.1e-165;
Matches 458; Conservative 44; Mismatches 54; Indels 2; Gaps 2;
QY 1 MSKLRRVVTVLALSMASFFASGKIEAAAAESLATRTASTENSDNDNVFQATAKKVREG 59
Db 1 MSKLRRVVTVLALSMASFFASGKIEAAAAESLATRTASTENSDNDNVFQATAKKVREG 60
QY 60 RRNKQVEQKSRGAFCDKFEFPCCEGRGCPVEAQOESCYGRLYSVKVNDDCNVEICQSP 118
Db 61 RRNKQVEQKHTGAFCDKFEFPCGGGOCQPDVATQESCYGKMYCVRVNDDCNVEISQSP 119
QY 120 EVATVGSPPYPIELAIAGKKDCVDVVTITQOLPCEAEFVSSDPETPTSDGKLWVKIDRLGA 178
Db 121 EVATVGSPPYPIELAVGKDCVNVVITQOLPCEVEFVSSDPATPTSDSKLIWTIDRLGQ 180
QY 180 GDCKKITVWVKPLKEGCCFTAAATVCACPELRSYTKGQPAICIKQEGPCDCAELRCPCVYK 238
Db 181 GECKKITVWVKPLKEGCCFTAAATVCACPELRSYTKGQPAICIKQEGPECAELRCPCVYK 240
QY 240 IEVNTGSAIARNVTVNDNPVPGYSHASQORVLSFNLGDMRPGDKVFTVEFCPORRGO 298
Db 241 IEVNTGSAIARNVTVNDNPVPGYSHASQORVLSFNLGDMRPGDKVFTVEFCPORRGO 300
QY 300 TNVATVYCGGHKCSANVTTVNNEPCQVNVISGADWSYVCKPVEYSISVSNPGDLVLHD 358
Db 301 TNVATVYCGGHKCSANVTTVNNEPCQVNVISGADWSYVCKPVEYTIIVSNPGDLKLYD 360
QY 360 VIQDTLPSTGVTLEAPGGEICCNKVMRIKEMCPGETLQFKLVVKAQVPGFTNOVAVT 418
Db 361 VIEDTAPSGATILEAAGAEICCNKAVMCIKEMCPGETLQFKLVVKAQVPGFTNOVAVT 420
QY 420 ESNCGTCTSCAETHHMKGLAATHMCLVDTNDPVCVGENTVYRICVTVNRGSAEDTNVSL 478
Db 421 NSDCGTCTSCAETHHMKGLAATHMCLVDTNDPVCVGENTVYRICVTVNRGSAEDTNVSL 480
QY 479 ILKFSKELOPIASSGPTKGTISGNTVVFDPALPKLGSKESEFVSITLKGAPDARGEAIL 538
Db 481 ILKFSKELOPVSSSGPTKGTITGNTVVFDPALPKLGSKESEFVSITLKGAPDARGEAIL 540
QY 539 SSDDLTPSPVSDTENTHVV 556
Db 541 SSDDLTPVPVADTENTHVV 558
RESULT 5
C81671
60 kDa outer membrane protein TC0727 [imported] - Chlamydia muridarum (strain Nigg)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000
C:Accession: C81671
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
/ C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A:Reference number: A81500; MUID:20150255
A:Accession: C81671
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-554 <TET>
A:Cross-references: GB:AE002341; GB:AF002160; NID:g7190754; PIDN:AAF39537.1; PID:g71
A:Experimental source: strain Nigg (MoPn)
C:Genetics:
A:Gene: TC0727
C:Superfamily: 60K cysteine-rich outer membrane protein
Query Match 73.7%; Score 2176; DB 2; Length 554;
Best Local Similarity 72.8%; Pred. No. 1.8e-143;
Matches 405; Conservative 74; Mismatches 69; Indels 8; Gaps 3;
QY 1 MSKLRRVVTVLALSMASFFASGKIEAAAAESLATRTASTENSDNDNVFQATAKKVREG 60
Db 7 MNKLIRRAVTVFAVTSVASFASGVLETSMAESLSTNVLADTK-AKETTSQDKRKAR 65
QY 61 RRNKQVEQKSRGAFCDKFEFPCCEGRGCPVEAQOESCYGRLYSVKVNDDCNVEICQSP 120
Db 66 KKHQ-----NRTSVYRKEVTAVRDTKA--VEPRQSCFCGKMYTVKVNDDNRNVEIVQSVPE 118
QY 121 YATVGSPPYPIELAIAGKKDCVDVVTITQOLPCEAEFVSSDPETPTSDGKLWVKIDRLGAG 180
Db 119 YATVGSPPYPIELAIAGKKDCVDVVTITQOLPCEAEFVSSDPATPTADGKLWVKIDRLGOG 178

Db 1 MSKLRRVVTVLALSMASFFASGKIEAAAAESLATRTASTENSDNDNVFQATAKKVREG 60
QY 60 RRNKQVEQKSRGAFCDKFEFPCCEGRGCPVEAQOESCYGRLYSVKVNDDCNVEICQSP 119
Db 61 RRNKQVEQKHTGAFCDKFEFPCGGGOCQPDVATQESCYGKMYCVRVNDDCNVEISQSP 120
QY 120 EVATVGSPPYPIELAIAGKKDCVDVVTITQOLPCEAEFVSSDPETPTSDGKLWVKIDRLGA 179
Db 121 EVATVGSPPYPIELAVGKDCVNVVITQOLPCEVEFVSSDPATPTSDSKLIWTIDRLGQ 180
QY 180 GDCKKITVWVKPLKEGCCFTAAATVCACPELRSYTKGQPAICIKQEGPCDCAELRCPCVYK 239
Db 181 GECKKITVWVKPLKEGCCFTAAATVCACPELRSYTKGQPAICIKQEGPECAELRCPCVYK 240
QY 240 IEVNTGSAIARNVTVNDNPVPGYSHASQORVLSFNLGDMRPGDKVFTVEFCPORRGOI 299
Db 241 IEVNTGSAIARNVTVNDNPVPGYSHASQORVLSFNLGDMRPGDKVFTVEFCPORRGOI 300
QY 300 TNVATVYCGGHKCSANVTTVNNEPCQVNVISGADWSYVCKPVEYSISVSNPGDLVLHDV 359
Db 301 TNVATVYCGGHKCSANVTTVNNEPCQVNVISGADWSYVCKPVEYTIIVSNPGDLKLYD 360
QY 360 VIQDTLPSTGVTLEAPGGEICCNKVMRIKEMCPGETLQFKLVVKAQVPGFTNOVAVTS 419
Db 361 VIEDTAPSGATILEAAGAEICCNKAVMCIKEMCPGETLQFKLVVKAQVPGFTNOVAVTS 420
QY 420 ESNCGTCTSCAETHHMKGLAATHMCLVDTNDPVCVGENTVYRICVTVNRGSAEDTNVSLI 479
Db 421 NSDCGTCTSCAETHHMKGLAATHMCLVDTNDPVCVGENTVYRICVTVNRGSAEDTNVSLI 480
QY 480 LKFSKELOPIASSGPTKGTISGNTVVFDPALPKLGSKESEFVSITLKGAPDARGEAIL 539
Db 481 LKFSKELOPVSSSGPTKGTITGNTVVFDPALPKLGSKESEFVSITLKGAPDARGEAIL 540
QY 540 SDTLTSPVSDTENTHVV 556
Db 541 SDTLTVPVADTENTHVV 557
RESULT 4
JC5204
60K cysteine-rich outer membrane protein 2 precursor - Chlamydomophila psittaci
C:Species: Chlamydomophila psittaci, Chlamydia psittaci
C:Date: 20-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 31-Mar-2000
C:Accession: JC5204
R:Hisia, R.; Bavoil, P.M.
Gene 176, 155-162, 1996
A:Title: Sequence analysis of the omp2 region of Chlamydia psittaci strain GPIC: Structu
A:Reference number: JC5203; MUID:97075924
A:Accession: JC5204
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-558 <HSI>
A:Cross-references: GB:U41759; NID:g1783376; PIDN:AAB41143.1; PID:g1783382
A:Experimental source: strain GPIC
C:Genetics:
A:Gene: omp2
C:Superfamily: 60K cysteine-rich outer membrane protein
C:Keywords: membrane protein; virulence
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-40/Domain: propeptide #status predicted <PRO>
F:41-558/Product: 60K cysteine-rich outer membrane protein 2 #status predicted <MAT>
Query Match 84.38; Score 2488; DB 2; Length 558;
Best Local Similarity 82.1%; Pred. No. 4.1e-165;
Matches 458; Conservative 44; Mismatches 54; Indels 2; Gaps 2;
QY 1 MSKLRRVVTVLALSMASFFASGKIEAAAAESLATRTASTENSDNDNVFQATAKKVREG 59
Db 1 MSKLRRVVTVLALSMASFFASGKIEAAAAESLATRTASTENSDNDNVFQATAKKVREG 60
QY 60 RRNKQVEQKSRGAFCDKFEFPCCEGRGCPVEAQOESCYGRLYSVKVNDDCNVEICQSP 118
Db 61 RRNKQVEQKHTGAFCDKFEFPCGGGOCQPDVATQESCYGKMYCVRVNDDCNVEISQSP 119
QY 120 EVATVGSPPYPIELAIAGKKDCVDVVTITQOLPCEAEFVSSDPETPTSDGKLWVKIDRLGA 178
Db 121 EVATVGSPPYPIELAVGKDCVNVVITQOLPCEVEFVSSDPATPTSDSKLIWTIDRLGQ 180
QY 180 GDCKKITVWVKPLKEGCCFTAAATVCACPELRSYTKGQPAICIKQEGPCDCAELRCPCVYK 238
Db 181 GECKKITVWVKPLKEGCCFTAAATVCACPELRSYTKGQPAICIKQEGPECAELRCPCVYK 240
QY 240 IEVNTGSAIARNVTVNDNPVPGYSHASQORVLSFNLGDMRPGDKVFTVEFCPORRGO 298
Db 241 IEVNTGSAIARNVTVNDNPVPGYSHASQORVLSFNLGDMRPGDKVFTVEFCPORRGO 300
QY 300 TNVATVYCGGHKCSANVTTVNNEPCQVNVISGADWSYVCKPVEYSISVSNPGDLVLHD 358
Db 301 TNVATVYCGGHKCSANVTTVNNEPCQVNVISGADWSYVCKPVEYTIIVSNPGDLKLYD 360
QY 360 VIQDTLPSTGVTLEAPGGEICCNKVMRIKEMCPGETLQFKLVVKAQVPGFTNOVAVT 418
Db 361 VIEDTAPSGATILEAAGAEICCNKAVMCIKEMCPGETLQFKLVVKAQVPGFTNOVAVT 420
QY 420 ESNCGTCTSCAETHHMKGLAATHMCLVDTNDPVCVGENTVYRICVTVNRGSAEDTNVSL 478
Db 421 NSDCGTCTSCAETHHMKGLAATHMCLVDTNDPVCVGENTVYRICVTVNRGSAEDTNVSL 480
QY 479 ILKFSKELOPIASSGPTKGTISGNTVVFDPALPKLGSKESEFVSITLKGAPDARGEAIL 538
Db 481 ILKFSKELOPVSSSGPTKGTITGNTVVFDPALPKLGSKESEFVSITLKGAPDARGEAIL 540
QY 539 SSDDLTPSPVSDTENTHVV 556
Db 541 SSDDLTPVPVADTENTHVV 558
RESULT 5
C81671
60 kDa outer membrane protein TC0727 [imported] - Chlamydia muridarum (strain Nigg)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000
C:Accession: C81671
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
/ C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A:Reference number: A81500; MUID:20150255
A:Accession: C81671
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-554 <TET>
A:Cross-references: GB:AE002341; GB:AF002160; NID:g7190754; PIDN:AAF39537.1; PID:g71
A:Experimental source: strain Nigg (MoPn)
C:Genetics:
A:Gene: TC0727
C:Superfamily: 60K cysteine-rich outer membrane protein

```
QY 181 DKCKITVWVKPLKEGCGCFTAAATYACAPELRSYTKGQPAICIKQEGPDCAICLRCPVCYKI 240
Db 179 ESKITVWVKPLKEGCGCFTAAATYACAPELRSYTKGQPAICIKQEGPDCAICLRCPVYRI 238
QY 241 EYVNTGSATARNVTVDNVPDGYSHASGQRLVFNIGDMRPGDKKVFTEFCFQRRGQIT 300
Db 239 NVVNOGTATARNVVNPNVDPDGYAHASGQRLVFTYLGDMQPGQRITITVEFCPLKGRVT 298
QY 301 NVATVYCGGHKCSANVTTVVNEPCVQVNIISGADWSYVCKPVEYSISVSNSPGDLVLHDV 360
Db 299 NIATVSYCGGHKNTASVTTVINEPCVQVNIISGADWSYVCKPVEYISVSNSPGDLVLHDV 358
QY 361 IODTLPSGVTVLEAPGEGELCCNKKVVRKEMCPGETLQFKLVYKAQVPGRTNOVAVTSE 420
Db 359 IEDTLSPGLTVVEAAGAQISCNKLVWTLKELNPGESLOYKVLVRAQTPOGTNNVVVKSC 418
QY 421 SNGCTCTCAEATTHWKGIAATHMCLVDLNDPICVGENTVYRICVNTNRGSAEDTNVSLIL 480
Db 419 SDGICTSCAEATTVWKGVAATHMCLVDLNDPICVGENTVYRICVNTNRGSAEDTNVSLIL 478
QY 481 KFSKELQIPASSGPTKGIISGNTVVDALPKLGSKESVEFSVTLKGIAPGDARGEAILSS 540
Db 479 KFSKELQIPASSGPTKGIISGNTVVDALPKLGSKESVEFSVTLKGIAPGDARGEAILSS 538
QY 541 DTLTSPVSDTENTHY 556
Db 539 DTLTVPVSDTENTHIY 554

RESULT 6
B43584
60K cysteine-rich outer membrane protein 2 precursor, serotype E - Chlamydia trachomatis
C:Species: Chlamydia trachomatis
C>Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 20-Aug-1999
C:Accession: B43584; S13120; S18979; S24275
R:de la Maza, L.M.; Fielder, T.J.; Carlson, E.J.; Markoff, B.A.; Peterson, E.M.
Nucleic Acids Res. 18, 6713, 1990
A:Title: The nucleotide and derived amino acid sequence of the omp2 gene of Chlamydia trachomatis
A:Reference number: S13120; MUID:91067486
A:Accession: S13120
A:Molecule type: DNA
A:Residues: 1-547 <DEL>
A:Cross-references: EMBL:X54389; NID:g40763; PIDN:CAA38259.1; PID:g40764
A:Experimental source: strain Bour, serotype E
R:Coles, A.M.; Allan, I.; Pearce, J.H.
Nucleic Acids Res. 18, 6713, 1990
A:Title: The nucleotide and derived amino acid sequence of the omp2 gene of Chlamydia trachomatis
A:Reference number: S13120; MUID:91067486
A:Accession: S13120
A:Molecule type: DNA
A:Residues: 1-32, 'FT', 35-120, 'L', 122-131, 'A', 133-457, 'S', 459-547 <COL>
A:Cross-references: EMBL:X55903; NID:g40724; PIDN:CAA39396.1; PID:g40725
A:Experimental source: strain DK20, serotype E
C:Genetics:
A:Gene: omp2
A:Function:
A:Description: associated with differentiation of reticulate bodies into elementary bodies
A:Note: essential for the structural integrity of the outer envelope of the elementary body
A:Superfamily: 60K cysteine-rich outer membrane protein
C:Keywords: membrane protein; virulence
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-40/Domain: propeptide #status predicted <PRO>
F:41-547/Product: 60K cysteine-rich outer membrane protein 2 #status predicted <MAT>

Query Match 73.3%; Score 2162.5; DB 2; Length 547;
Best Local Similarity 72.4%; Pred. No. 1.6e-142;
Matches 403; Conservative 72; Mismatches 71; Indels 11; Gaps 4;

QY 1 MSLKIRRVTVLALTSMAISCFASGIEAAVAESLTKIVASAEETPAPVMTAKKVLVR 60
Db 1 MNKLIRRAVTIFAVTSVASLFSAGVLETSMAESLSTNVISLADTKAK--DNTSHKSKKAR 58
```

```
QY 61 RNQOPVEQKSRGAFCD--KEFYPCEEGRCPVEAQOESCGRGLYSVKVNDNCNVEICOSVP 119
Db 59 KN-----HSKETLVDRKEVAPVHESKA--TGRQDSCFGRMYTVKVNDNRNVEITOAVP 110
QY 120 EYATVSGPYPIETLATGKDCVDVVIITQOLPCEAEFVSSDPETPTSDGKLWVKIDRLGA 179
Db 111 EYATVSGPYPIETLATGKDCVDVVIITQOLPCEAEFVSSDPETPTSDGKLWVKIDRLGA 170
QY 180 GDRCKITVWVKPLKEGCGCFTAAATYACAPELRSYTKGQPAICIKQEGPDCAICLRCPVCYK 239
Db 171 GEKSKITVWVKPLKEGCGCFTAAATYACAPELRSYTKGQPAICIKQEGPDCAICLRCPVCYK 230
QY 240 IEVWNTGSATARNVTVDNVPDGYSHASGQRLVFNIGDMRPGDKKVFTEFCFQRRGQI 299
Db 231 INVWNTGSATARNVTVDNVPDGYSHASGQRLVFNIGDMRPGDKKVFTEFCFQRRGQI 290
QY 300 TNVATVYCGGHKCSANVTTVVNEPCVQVNIISGADWSYVCKPVEYSISVSNSPGDLVLHDV 359
Db 291 TNVATVYCGGHKNTASVTTVINEPCVQVNIISGADWSYVCKPVEYISVSNSPGDLVLHDV 350
QY 360 VIQDPLSPGVTVLEAPGEGELCCNKKVVRKEMCPGETLQFKLVYKAQVPGRTNOVAVTSE 419
Db 351 VVEDTLSPGVTVLEAPGEGELCCNKKVVRKEMCPGETLQFKLVYKAQVPGRTNOVAVTSE 410
QY 420 ESNCGTCTCAEATTHWKGIAATHMCLVDLNDPICVGENTVYRICVNTNRGSAEDTNVSLI 479
Db 411 CSDGCTCTCAEATTHWKGIAATHMCLVDLNDPICVGENTVYRICVNTNRGSAEDTNVSLI 470
QY 480 LKFSKELQIPASSGPTKGIISGNTVVDALPKLGSKESVEFSVTLKGIAPGDARGEAILSS 539
Db 471 LKFSKELQIPASSGPTKGIISGNTVVDALPKLGSKESVEFSVTLKGIAPGDARGEAILSS 530
QY 540 DTLTSPVSDTENTHY 556
Db 531 DTLTVPVSDTENTHIY 547

RESULT 7
60K cysteine-rich outer membrane protein 2 precursor serotypes B, C, and D - Chlamydia trachomatis
C:Species: Chlamydia trachomatis
C>Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 20-Aug-1999
C:Accession: D71515; C43584; S11673
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis
A:Reference number: A71570; MUID:99000809
A:Accession: D71515
A:Molecule type: DNA
A:Residues: 1-553 <ARN>
A:Cross-references: GB:AE001317; GB:AE001273; NID:g3328863; PIDN:AA68042.1; PID:g332
A:Experimental source: serotype D, strain UW-3/Cx
R:de la Maza, L.M.; Fielder, T.J.; Carlson, E.J.; Markoff, B.A.; Peterson, E.M.
Infect. Immun. 59, 1196-1201, 1991
A:Title: Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodal
A:Reference number: A43584; MUID:91147205
A:Accession: C43584
A:Molecule type: DNA
A:Residues: 7-238, 'V', 240-553 <DEL>
A:Cross-references: GB:X54388; NID:g40760; PIDN:CAA38257.1; PID:g40761
A:Experimental source: serotype C
R:Watson, M.W.; Lambden, P.R.; Ward, M.E.; Clarke, I.N.
FEMS Microbiol. Lett. 65, 293-297, 1989
A:Title: Chlamydia trachomatis 60 kDa cysteine rich outer membrane protein: sequence
A:Reference number: S11673
A:Accession: S11673
A:Molecule type: DNA
A:Residues: 7-553 <WAT>
A:Cross-references: EMBL:X53510; NID:g40681; PIDN:CAA37588.1; PID:g40683
A:Experimental source: serotype B
C:Genetics:
A:Gene: omp2; omcB
```


Tue May 28 08:55:19 2002

Q:Function:
A:Description: associated with differentiation of reticulate bodies into elementary bodies
A:Note: essential for the structural integrity of the outer envelope of the elementary body
Q:Superfamily: 60K cysteine-rich outer membrane protein
Q:Keywords: membrane protein; virulence
P:1-28/Domain: signal sequence #status predicted <SIG>
P:29-46/Domain: propeptide #status predicted <PRO>
P:47-553/Product: 60K cysteine-rich outer membrane protein 2 #status predicted <MAT>

Query Match 73.2%; Score 2160.5; DB 2; Length 553;
Best Local Similarity 71.7%; Pred. No. 2.2e-142; Indels 15; Gaps 4;
Matches 401; Conservative 75; Mismatches 68;

QY 1 MSKLIRRVTVLALTSMAFCGAEAEVSLITKIVASAEKPAVPMTAKKVLVR 60
DB 7 MNKLIRRAVTIFAVTSVASLFSAGVLETSMAESLSTNVISLADTKAK--DNTSHKSKKAR 64
QY 61 RN---KQPVQKSRGAFCDKEFYPCCEGRQCPVEAQOQSCYGRLYSVKYNDDCNVEICQS 117
DB 65 KHSKETPVDR-----KEVAPVHESKA--TGPQKQSCFGRMVTVKVNDRNVEITQA 114
QY 118 VPEYATVGSYPYIEILAIKCKDCVDVITQOLPCEAEFVSDDPETTPTSDGKLVWKIDRL 177
DB 115 VPEYATVGSYPYIEITATGRKDCVDVITQOLPCEAEFVRSDDPATPTADGKLVWKIDRL 174
QY 178 GAGDKCKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKOEGPDCACLRCPVC 237
DB 175 GQGEKSKITVWVKPLKEGCGCFTAAATVCACPEIRSVTKCGQPAICVKQOEGPENACLRCPVV 234
QY 238 YKIEVNTVGSATARNVTVDNVPDGYSHASQORVLSFNGLDMRPGDKKVTVEFCPPORRG 297
DB 235 YKINIVNQGTATARNVVENPVDPGYAHSSGQRVLTFTLGDMPQGEHRTITVEFCPLKRG 294
QY 298 QITNVATVTCGGHKCSANVTTVVNEPCVQVNI SGADMSYVCKPVEYSISVSNPGDLVLH 357
DB 295 RATNIAMVSYCGGHKNFTASVTTVINEPCVQVSIAGADWSYVCKPVEYSISVSNPGDLVLR 354
QY 358 DWIQTDLPSGVTVLEAPGGEICCNKVVWRKEMCPGETLOFKLVVKAQVGRNTQNAV 417
DB 355 DVVVEDTLSPGVTVLEAAGAQISCNKVVWTKELNPGESLQYKVLVRAQTGQGTNNVVV 414
QY 418 TSENCGCTCAETTHWKGLAATHMCVLDNDPICVGENTVYRICVNRGSAEDTNVS 477
DB 415 KSCSDCGCTCAETTHWKGVAATHMCVLDNDPICVGENTVYRICVNRGSAEDTNVS 474
QY 478 LILKFSKELQIASGPTKGTISGNVTVFDALPKLGSKESVEFSVTLKGIAPGDARGEAI 537
DB 475 LMLKFSKELQVPVSFGPTKGTITGNTVTVFDLSPLRGLSKETVEFSVTLKAVSAGDARGEAI 534
QY 538 LSSDTLTSFVSDTENTHYV 556
DB 535 LSSDTLTVFVSDTENTHIY 553

RESULT 8
A32244
60K cysteine-rich outer membrane protein 2 precursor, serotype L1 and L2 - Chlamydia tra
C:Species: Chlamydia trachomatis
C:Date: 12-Oct-1989 #sequence_revision 27-Jun-1994 #text_change 16-Jul-1999
C:Accession: A32244; A33584; A36043; A30472; JT0419; S18981; S24277
J:Allen, J.E.; Stephens, R.S.
J: Bacteriol. 171, 285-291, 1989
A:Title: Identification by sequence analysis of two-site posttranslational processing of
A:Reference number: A32244; MUID:89123030
A:Accession: A32244
A:Molecule type: DNA
A:Residues: 1-547 <ALL>
A:Cross-references: GB:M23001; NID:g144552; PIDN:AAA23152.1; PID:g144553
A:Experimental source: strain L2/434/Bu
A:Note: parts of this sequence, including the amino ends of the precursor and mature pro
R:De la Maza, L.M.; Fielder, T.J.; Carlson, E.J.; Markoff, B.A.; Peterson, E.M.
Infect. Immun. 59, 1196-1201, 1991

A:Title: Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodal
A:Reference number: A43584; MUID:91147205
A:Accession: A43584
A:Molecule type: DNA
A:Residues: 1-547
A:Cross-references: GB:M23001; NID:g144552; PIDN:AAA23152.1; PID:g144553
A:Experimental source: serovar 2, strain L2/434/Bu
R:Wahlberg, J.; Lundberg, J.; Hultman, T.; Uhlen, M.
Proc. Natl. Acad. Sci. U.S.A. 87, 6569-6573, 1990
A:Title: General colorimetric method for DNA diagnostics allowing direct solid-phase
A:Reference number: A36043; MUID:90370827
A:Accession: A36043
A:Molecule type: DNA
A:Residues: 294-402 <WAH>
A:Experimental source: serotype L2
R:Lambden, P.R.; Everson, J.S.; Ward, M.E.; Clarke, I.N.
Gene 87, 105-112, 1990
A:Title: Sulfur-rich proteins of Chlamydia trachomatis: developmentally regulated tra
A:Reference number: JQ0514; MUID:90236284
A:Accession: A30472
A:Molecule type: DNA
A:Residues: 1-46; 528-547 <LAM>
A:Cross-references: GB:M35148; GB:M23180; GB:M35161; NID:g144485
R:Clarke, I.N.; Ward, M.E.; Lambden, P.R.
Gene 71, 307-314, 1988
A:Title: Molecular cloning and sequence analysis of a developmentally regulated cyste
A:Reference number: JT0419; MUID:89138006
A:Accession: JT0419
A:Molecule type: DNA
A:Residues: 30-547 <CLA>
A:Cross-references: GB:M35148; NID:g144485; PIDN:AAA23119.1; PID:g144487
A:Experimental source: serotype L1
C:Genetics:
A:Gene: omp2; omcB
C:Function:
A:Description: associated with differentiation of reticulate bodies into elementary b
A:Note: essential for the structural integrity of the outer envelope of the elementar
C:Superfamily: 60K cysteine-rich outer membrane protein
C:Keywords: membrane protein; virulence
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-40/Domain: propeptide #status experimental <PRO>
F:41-547/Product: 60K cysteine-rich outer membrane protein 2 #status experimental <MA

Query Match 72.2%; Score 2129.5; DB 1; Length 547;
Best Local Similarity 70.7%; Pred. No. 3e-140;
Matches 395; Conservative 77; Mismatches 72; Indels 15; Gaps 4;

QY 1 MSKLIRRVTVLALTSMAFCGAEAEVSLITKIVASAEKPAVPMTAKKVLVR 60
DB 1 MNKLIRRAVTIFAVTSVASLFSAGVLETSMAEFISLADTKAK--DNTSHKSKKAR 58
QY 61 RN---KQPVQKSRGAFCDKEFYPCCEGRQCPVEAQOQSCYGRLYSVKYNDDCNVEICQS 117
DB 59 KHSKETPVNR-----KKVAPVHESKA--TGPQKQSCFGRMVTVKVNDRNVEITQA 108
QY 118 VPEYATVGSYPYIEILAIKCKDCVDVITQOLPCEAEFVSDDPETTPTSDGKLVWKIDRL 177
DB 109 VPEYATVGSYPYIEITATGRKDCVDVITQOLPCEAEFVRSDDPATPTADGKLVWKIDRL 168
QY 178 GAGDKCKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKOEGPDCACLRCPVC 237
DB 169 GQGEKSKITVWVKPLKEGCGCFTAAATVCACPEIRSVTKCGQPAICVKQOEGPENACLRCPVV 228
QY 238 YKIEVNTVGSATARNVTVDNVPDGYSHASQORVLSFNGLDMRPGDKKVTVEFCPPORRG 297
DB 229 YKINIVNQGTATARNVVENPVDPGYAHSSGQRVLTFTLGDMPQGEHRTITVEFCPLKRG 288
QY 298 QITNVATVTCGGHKCSANVTTVVNEPCVQVNI SGADMSYVCKPVEYSISVSNPGDLVLH 357
DB 289 RATNIAMVSYCGGHKNFTASVTTVINEPCVQVSIAGADMSYVCKPVEYSISVSNPGDLVLR 348

Qy 358 DVVIQDTLSPGVTVLEAPGGEICCNKVMIRIKEMCPGETLQFKLVKVAQVPGRETNOVAV 417
Db 349 DVVVKDTLSPGVTVLEAAGAGACSCNKVYTKVKNLPGESLQKVLVRAQTGQCFNNVVV 408
Qy 418 TSESCGCTCAEATTTTHKGLAATHMCLVLTNDPICVGTNTYVYICVNTNRGSAEDTNS 477
Db 409 KSCDCGCTCAEATTTTHKGLAATHMCLVLTNDPICVGTNTYVYICVNTNRGSAEDTNS 468
Qy 478 LILFKELOPIASSGPTKGTISGNTVVFDPALPKLGSKESVEFSVTLKGIAPGDARGEAI 537
Db 469 LMLKFSKELQPVSPGPTKGTITGNTVVFDSLPRLGSKETVEFSVTLKAVSAGDARGEAI 528
Qy 538 LSSDTLSPVSDTENTHY 556
Db 529 LSSDTLTPVSDTENTHY 547

RESULT 9
F69009
probable membrane protein MTH1074 - Methanobacterium thermoautotrophicum (strain Delta H
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: F69009
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A:Reference number: A69000; MUID:98037514
A:Accession: F69009
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1474 <MTH>
A:Cross-references: GB:AE000878; GB:AE000666; NID:g2622171; PID:AAB85563.1; PID:g262217
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1074
C:Keywords: duplication

Query Match 5.1%; Score 151.5; DB 2; Length 1474;
Best Local Similarity 24.3%; Pred. No. 0.019;
Matches 83; Conservative 48; Mismatches 159; Indels 51; Gaps 10;
Qy 236 VCYKIEVNTGSAIARNVTVDPDGYSHASQORVLSFNLGDMRGDKKVTVEFCPOR 295
Db 382 VRFTITVNVGNPNATGVVYVTDLLPQLSFVSAS--ASRGITNTTGTWIGNLEYFETV 439
Qy 296 RGOITNAVITYCGHKCSANTVTVVNEPCVQVNISGA-----DWS-----Y 337
Db 440 TLNIT--ATVTATGATVNNANTVGDVDPDMANNVSAALNPPASDLTIDKSVNPNPEY 497
Qy 338 VCKPVEYSISVSNPGLVLDHVVIODTLPSGVTVLEAPGGEICCNKVMIRIKEMCPGETL 397
Db 498 VGENIQYITVSNRCPDNaGVVDEVLPAGLIPISATPSKGSYNGTNNVGLNYLEIA 557
Qy 398 QFKLVKVAQVPGRETNOVAVTS---ESNCGTCTCAETTTTHKGL--AATHMCLVLTNDP- 452
Db 558 TLTIARVNTAGSLTNFANITSPNPDNPNNDPAEVV---GIPVADLLIVKQVSDPR 613
Qy 453 ICVGTNVYICVNTNRGSAEDTNSLILKFSKELQPIASSGPTKGTISGNTVVFDPALPKL 512
Db 614 PDYGSVTVTVAVNLGPNATGVTVTDILSPGLVYL-SHTVTOGTNATTGVN----- 666
Qy 513 GSKESVEFSVTLKGIAPDARGEAILSSDTLTPSPVSDTENT 553
Db 667 -----YIGALNYAASALMNLTVLVNTTGDNSNT 694

RESULT 10
T28669
surface protein 51c - Paramesium tetraurelia
C:Species: Paramesium tetraurelia

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 07-Dec-1999
C:Accession: T28669
R:Nielsen, E.; You, Y.; Forney, J.
J. Mol. Biol. 222, 835-841, 1991
A:Title: Cysteine residue periodicity is a conserved structural feature of variable s
A:Reference number: 220504; MUID:92106337
A:Accession: T28669
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-2233 <NIE>
A:Cross-references: EMBL:M65164; NID:g159974; PID:g159975; PIDN:AAA61740.1
C:Genetics:
A:Genetic code: SOCS

Query Match 4.6%; Score 135; DB 2; Length 2233;
Best Local Similarity 20.0%; Pred. No. 0.41; Mismatches 71; Indels 200; Gaps 34;
Matches 125; Conservative 71; Mismatches 71; Indels 200; Gaps 34;
Qy 5 IRRVVTVLALTSMAFCASGAGIEAAVAESLI-----TKIVASAFETKPAV 49
Db 1 MKRTLLIAMIATATCOVVSKEACTCAQLTSGDCARNSNCNNTTKLACEVQSTGVP 60
Qy 50 PMTAKKVLVRRNKQPVQKSRGAFCDKEFYPCBEGRCOPVEAOQESCYG----- 99
Db 61 TVT-----KNYKSLYC-----EGLAQTDCLKLNCAWIDNKTCTFTS 98
Qy 100 RLYSVKVNDDC-----NVEICQSVPEVATVGSVPYIE-----LAIGKK 138
Db 99 CTPEYKTIKDCOAISKRCITDGTICVEIDLCT---YLTSTCYONKAGNYCVWDETAK 155
Qy 139 DCVDVVITQOLP-----CEA--EFVSDPETTPSDGKLVMKIDRLGAGDKCKITVM 188
Db 156 KCSDEVTECAQLTALTKDSECRALYKF-----ECTAKPAGGCV-----DSGTNCADQVS 204
Qy 189 VKPLKEGC-----CFTAATVCAPELRSYTKCOGPAICIKOEGPDCACLRCPVCYK 239
Db 205 V-----EGCVTKTRSVNCFWDTTNNKCFD---KKENASTNKTHT-VDCQAF-LPTC-- 252
Qy 240 IEVNVNTGSAIARNVTVDPDGYSHASQORVLSFNLGDMRGDK-----KVFTVEFC 292
Db 253 -TAKGGGCVDIKTCADGKIKECKIDSAKKECYWSDKDLCKDKIKASAPNTLTNNSDC 311
Qy 293 POR--RQGITNVATV---TYCGGHKCSANTVTVVNEPCV--OVNIGADWSYCKPVEYS 345
Db 312 QKQFLASCITNGAGCVDDTSCG-----SSVQEQCAVNRNNRECTWNGSCK----D 359
Qy 346 ISVSNPG-DLVLDH-----VVI-----QDTLPSPGV 370
Db 360 KTCENAGTDIVGHDOGSTYKACTGKANNAGGCKRSCDAPTTIISNAGCEDYLPNGKC 419
Qy 371 VLEAPGGEI---CCNKVWNR---IKEM---CPGETLQFKLVK--AQVPGRETNOVAV 418
Db 420 IAKKGGCISNTTCSAILLKDACVQDNKKQCYWDTVGNCLDKTCATLPTRLNLSHLCN 479
Qy 419 SESNGCTCT-----SCAETTTTHKGLAATHMCLVLTNDPICVGTNVYICVNTNRGSAED 473
Db 480 GEIN--TCTVSSSGTCVDLLC---ENVVDKNCVKDKSGADCV---YYGSCYQKQCSAAS 531
Qy 474 TNVSLILKFSKELQPIASSGPTKGTI 499
Db 532 QDNTTHAOQOEYLPACTLSNTKKGCI 557

RESULT 11
D75622
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: D75622
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: D75622
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-756 <WHI>
A:Cross-references: GB:AE001826; PID:96460827; PIDN:AAF12630.1; PID:96460926; TIGR:DRB00
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRB0039
A:Map position: megaplasmid
A:Genome: plasmid
A:Note: plasmid MPI

Query Match 4.4%; Score 130.5; DB 2; Length 756;
Best Local Similarity 22.6%; Pred. No. 0.25;
Matches 121; Conservative 68; Mismatches 217; Indels 129; Gaps 31;

QY 108 DDCNV-EICQSVPEVATGSPYPIEILAIKGGKDCVGVVIT-----QOL 149
DB 23 DDNNVAQLVATPPHLELRKAFGAERLRPGDGTGVVTTATNTGAGAAPDVVLEDPGLGOOL 82
QY 150 PCEAEFVSSDPPTTSDCKLWKID-----RLGAGDKCKIT--V 187
DB 83 DAGLAV---PGSARTDRGLEYSADGASQAAEPAAVRGVRVQAGTLEPGEQTTLTFRM 139
QY 188 WKPLKEGCGF-TAATVCACP-----ELR-----SYTKGCPAICIKOEGPDC-- 229
DB 140 EAPPSAENLLNNVATSVSVTGEQAQSDTLQVFLPGVAGLPQGP---LAPECTAADG 196
QY 230 -----ACLRCPVCYKTEVNTGSA-----IARNVTVDNPPVPGYSHASQORVLSNLGDM 279
DB 197 QTLVAVTEGPACFDHTVQNTGDMERLSVSTGQATPQLLGAAGEPLFPQPTLA-- 254
QY 280 RPDGKKVTFEFCORRGQITNVATVYTCGGHKCSANVT-----IWNNEPCVQVNIS-- 331
DB 255 -PGEOR--QVRVCYDLRSQAQPLAQV-IAOGERGTSNATDLRRVETQRPGLRKTVSKV 310
QY 332 GA-DW---SYVCK--PVEYSISVSNPGDLVLHDVLIQDTLPSPGVTVLEAPG-----GEIC 380
DB 311 GAPDWAAGSAVTSDELEYLTSLVNTNPAQPLAGVQVLDPLPAGTEFVSASDGGALLGAAS 370
QY 381 CNKYVWRKEMCPGETLQFLVKAQVPGR-----FTNOAVTSESNCGTCTSCAETT 433
DB 371 AAQVANTLGLDLPAGATRTLTLRVRV---GRVDRDQELRNWFELTSSLELPAPLHNAASA 427
QY 434 THWGLAATHMCVLDTDNDPICVGVNTVYRICVTNRGSAEDTNVSLI-----LKF--- 482
DB 428 VVW-NTAPLLSKTLDRD-AAPGDLTYTLTKN-PSASTALVDLVITDTPAAALKYVTG 484
QY 483 SKELOPIASSGPTKTISGNTVWFDPALPKLSKESVEFVTLKGIAPGDARGEAI 537
DB 485 TSLRAGVPTADPAE---NGGELQW-RVPRLGAGESLTLSYGLR-VLPG-AGELL 533

RESULT 12
S16148
gene serrate protein precursor - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 31-Dec-1991 #sequence_revision 02-Aug-1994 #text_change 17-Nov-2000
A:Accession: S16148; S16878; A36666
R:Thomas, U.; Speicher, S. A.; Knust, E.
A:Title: The Drosophila gene serrate encodes an EGF-like transmembrane protein with a co
Development 111, 749-761, 1991
A:Reference number: S16148; MUID:91347903
A:Accession: S16148
A:Molecule type: mRNA
A:Residues: 1-1408 <THOI>
A:Cross-references: EMBL:X56811
R:Thomas, U.
submitted to the EMBL Data Library, November 1990

A:Reference number: S16878
A:Accession: S16878
A:Molecule type: mRNA
A:Residues: 1-1351, 'T', 1353-1408 <THO2>
A:Cross-references: EMBL:X56811; NID:98563; PID:98564
R:Fleming, R. J.; Scottgale, T. N.; Diederich, R. J.; Artavanis-Tsakonas, S.
Genes Dev. 4, 2188-2201, 1990
A:Title: The gene Serrate encodes a putative EGF-like transmembrane protein essential
A:Reference number: A36666; MUID:91099666
A:Accession: A36666
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-15, 20-26, 'A', 28-1408 <FLE>
A:Cross-references: GB:M35759; NID:9158605; PID:9158606
C:Genetics:
A:Gene: FlyBase:Ser
A:Cross-references: FlyBase:FBgn0004197
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: glycoprotein; transmembrane protein
F:1-84/Domain: signal sequence #status predicted <SIG>
F:85-1408/Product: gene serrate protein #status predicted <MAT>
F:85-1221/Domain: extracellular #status predicted <EXT>
F:85-316/Domain: EGF homology <EG01>
F:319-348/Domain: EGF homology <EG02>
F:355-388/Domain: EGF homology <EG03>
F:395-488/Domain: EGF homology #status atypical <EG04>
F:495-526/Domain: EGF homology #status atypical <EG05>
F:533-608/Domain: EGF homology #status atypical <EG06>
F:615-645/Domain: EGF homology <EG07>
F:652-683/Domain: EGF homology <EG08>
F:690-720/Domain: EGF homology #status atypical <EG09>
F:727-796/Domain: EGF homology <EG10>
F:803-834/Domain: EGF homology <EG11>
F:841-876/Domain: EGF homology <EG12>
F:883-914/Domain: EGF homology <EG13>
F:921-952/Domain: EGF homology <EG14>
F:997-1060/Region: cysteine-rich
F:1222-1246/Domain: transmembrane #status predicted <TM1>
F:1247-1408/Domain: intracellular #status predicted <INT>
F:152,196,247,331,412,452,558,739,965,977,1004,1030,1150/Binding site: carbohydrate (

Query Match 4.4%; Score 130.5; DB 2; Length 1408;
Best Local Similarity 20.5%; Pred. No. 0.51;
Matches 124; Conservative 57; Mismatches 208; Indels 217; Gaps 32;

QY 14 LTSMSCFPASGGIEAAVAESLTIKIVASAEKTPAPVPMFAKRVLRNRKQVEOKSRGA 73
DB 547 LTTTATAITGSNLSSTALLAALTSVASTSLAIGPC-INAKECR-----NQP-----GS 594
QY 74 FCDKEFVPCBEGRCPVEAQEQSCYGRLYSVKVNDDCNVEICQSPVEYATVGSYPPIEL 133
DB 595 FA---CICKEG-----WGVTCENLDDC-VGQCRN----- 621
QY 134 AIGKKDCVDVVITQQLPCEAEFVSSDPE-----TPTSDG-----KLVMK-----IDRLG- 178
DB 622 ---GATCIDLVNDYRCACASGFTGRDCEIDIDECATSPCRNGGECVDMVGKFCICPLGY 678
QY 179 AGDKC---KITVWVKPLKEGCGCETA--ATVCAPELRSYTKCG-PAIC----- 221
DB 679 SGSLCEAEKENCPTSPCLGHLCLNTPEGYCHCPDRAGKHCEQLRPLCSOPPCNECCFA 738
QY 222 -----IKOEGPDC-----AKLRCPV---CYKIEVWNTGSAIA 250
DB 739 NVSLATSATTTTTTTTTTTTATTTTRKMAKPSGLPCSGHSGCEMSDVGTFC-KCHVHTGTGTC 797
QY 251 RNVTVNDVNP-----DG-----YSHASQGRVLS-----FNLGDMRPG-- 282
DB 798 HNLNECSPNCRNGGICLDGDFTCCECMGWTGKRCSEATGCGYAGOCQNGGTGMPGAP 857
QY 283 DKKV-----FTVEFCPQRRGQITNVAVTYCGGHKCSANVTTVVNEPCVQVNISGA 333
DB 858 DKALQPHCRCAPGWTGLFCAE-----AIDQCRGQFCHNGGT-----CE-----SGA 898

Qy	334	DW-SYVCKPVEYSISVNPGLDVLHWDVIQTDLPSGVTLEAPGGEICCNKVVVRIKEMC	392
Db	899	GWFRVCV-----AQGSGPDCRINNVNECSFQCQGATCIDGIGGYSC-----IC	943
Qy	393	PGETLQFKLVKAQVPGRTNQAVTSESNCGTCTCAET-----TTHWKLANATH	443
Db	944	P-----PGRHGRCEILLDPKSACQNASTSPYTALNRSONWLIDIALTG	989
Qy	444	MCVLDTNDPICGVNTVYRICVTVNGSABDTNWSILKFSELOPIASSGPFTGISCNT	503
Db	990	RTEDDENCAVCVN-----GTSRCTNLWGCLPNCYKVDPPLSKSSNLGSVCVKQHE	1039
Qy	504	VVFEDAL	509
Db	1040	VCVPAL	1045

RESULT 13
A48822
protein-glutamine gamma-glutamyltransferase (EC 2.3.2.13) - American bird grasshopper
N:Alternate names: annulin
C:Species: Schistocerca americana (American bird grasshopper)
C:Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 29-Oct-1999
C:Accession: A48822
R:Singer, M.A.; Hortsch, M.; Goodman, C.S.; Bentley, D.
Dev. Biol. 154, 143-159, 1992
A:Title: Annulin, a protein expressed at limb segment boundaries in the grasshopper embryo
A:Reference number: A48822; MUID:93050772
A:Accession: A48822
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-772 <SIN>
A:Cross-references: GB:M92291; NID:g160838; PID:g160839
A:Experimental source: embryo
A:Note: sequence extracted from NCBI backbone (NCBIN:117196, NCBIP:117199)
C:Superfamily: protein-glutamine gamma-glutamyltransferase
C:Keywords: aminocyclotransferase; blocked amino end; lipoprotein; myristylation; thioles
F2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F2/Modified site: aspartic acid (Asn) #status predicted
F4/5/Binding site: palmitate (Cys) (covalent) #status predicted
F4/341/Active site: Cys #status predicted

Query Match	4.4%	Score 129;	DB 2;	Length 772;	
Best Local Similarity	23.7%	Pred. No. 0.33;	118;	Indels	82; Gaps
Matches	79;	Conservative	55;	Mismatches	
Qy	41	SAETKPAVPMTAKVRLVRNK	-----QPVEQKSGRGAFCDKEF	-----YP	-----CCEGRGCP 89
Db	476	SGTQP	-----LKLIRKDLMLGIGNISTKAVGRQRE	DIINTYKY	PEKSVVEERAAM 526
Qy	90	VEA	---QQBSQGRLYSVKVNDDCN-VELQSVPEYATV	GPSYPTIILAIGK	KDCVDVVIT 146
Db	527	LKALRQSELSFRYY	---LNEFDNIHFNFELRDDIVIGSFVVVVV	MKNRNSQQDVTVT	583
Qy	147	QQLPCEAEFVSSDETTPTSDGKLVMKIDRL	---GAGDKCKITV	---	WKPLKEGCGFT 199
Db	584	VLLRDVTLV	-----TGHVKDGVKKEKVERLIKAGAVEE	RIDVSIEDYKY	KHLVDCAFN 638
Qy	200	AATVCACPELRSTYKCSQ	-----PAICIKQEGPDCACLRCPVCYKIE	VYNTGSAIARN	252
Db	639	IACLATVHD	-TNIYFAODDFVRVKPKDIIKLEGE	-----PV	-----QGQEMSAV 682
Qy	253	VTVDNPVP	-----DGYSHASQRV	-LSFNIGDMRPGDKKVTVEPC	PPORRGQITN 301
Db	683	ATLKNPLPIPVKKQGFLIEGPGIAKTQIKLSONIA	---PGEASVNFKFTPKYDGR	---	736
Qy	302	VATVTYCGGHKCSANVTTVN	-----EPCVQVNI	SG 332	
Db	737	-ATIAAKFSKELDDVDGFLNFMVEPKKEVNGTG	769		

RESULT 14
C75489

conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: F75489
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1
A:Reference number: A75250; MUID:20036895
A:Accession: F75489
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-998 <WHI>
A:Cross-references: GB:AE001925; GB:AE000513; NID:96458383; PIDN:AAF10262.1; PID:96458383
A:Experimental source: strain R1
A:Genetics:
A:Gene: DR0686
A:Map position: 1

Query Match	4.3%;	Score 127.5;	DB 2;	Length 998;
Best Local Similarity	21.9%;	Pred. No. 0.56;		
Matches 108;	Conservative 60;	Mismatches 181;	Indels 145;	Gaps 28;
QY	130	IEILIA-IGKKDCVDVITQQLPCEAEFVSSDPETTP-----TSDGKLVWKIDRLGAGDKC-	183	
Db	334	IEATAPICIRPVVEIALGPQ-----GNPRALPGGEMSAADRRQV-RDSVVGQGEVCF	383	
QY	184	-----KITVWVKPKLEGCCFTATVCAPELRSYTKGCPAICIKQEGPDCAC	231	
Db	384	THTVONIGLDUTDISTGVSRINEGANYTLRDMAGKINGDFVVTLEP-----QQTADFQAC	439	
QY	232	LROPVCYK---IEVYNT---GSAIARNVTVDN-----VPDGYSHAS---GQVLSFNLDGM	279	
Db	440	YSYPTAERTSLEAVLTSRSSOGAADNRVTDOITVPTNTGLTKTSDRGNAL-----V	493	
QY	280	RPDGKKVFTEFCPPQRGQITNATVTYCGGHKCSANVTTVNEPCVQVNVISGADWSYVC	339	
Db	494	APCEELTYTLTF-----TNTLPVPLNG-----VVVRDPLGVITVPDGTTP----	531	

RESULT 15

T25878
hypothetical protein T10E9.4 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T25878
R:Greco, T.; Hawkins, M.
submitted to the EMBL Data Library, April 1997
A:Description: The sequence of *C. elegans* cosmid T10E9.

Tue May 28 08:55:19 2002

A Reference number: Z20104
A Accession: T25878
A Status: preliminary; translated from GB/EMBL/DBJ
A Molecule type: DNA
A Residues: 1-693 <GRE>
A Cross-references: EMBL:U97403; PIDN:AAB52472.1; GSPDB:GN00019; CESP:T10E9.4
A Experimental source: strain Bristol N2; clone T10E9
C Genetics:
A Gene: CESP:T10E9.4
A Map position: 1
A Introns: 16/1; 40/1; 87/1; 136/1; 185/1; 264/1; 322/1; 488/1; 565/2

Query Match 4.3%; Score 127; DB 2; Length 693;
Best Local Similarity 19.8%; Pred. No. 0.4; Indels 222; Gaps 26;
Matches 103; Conservative 41; Mismatches 41; Indels 222; Gaps 26;

QY 46 PAPVMTAKKVLV-----RNKQFVEQK-----RGAFCDKKEFYPCEEG 85
DB 185 PIPFSFTLKNNSVPLTWAEGLKSKANNOSYKVKCTTNSCEGDGNGFCDVYNPTFKN 244
QY 86 -----RCQVEAQQESCYGR-----LYSVKV-----NDDC--NVEIC--QSVPEY 121
DB 245 HRFDRSGPKESLQFCYKQPDLPDLKHVKVDLGLKLCHYNEDCGNETEICHIDNIIVY 304
QY 122 ATVGSPPYPIEILAIKDKCDVVIQO-----LPCAEFVSSDPETPTSDGKLVM 172
DB 305 ANVTOPEKIAVPGI-----CVHVDVCSAEADDEKTHAVLPVNSQFC'KEDMHCQ----- 351
QY 173 KIDRLGAGDKCKIT-----VWVKPLKEGCECTAATVCAPELRSYTKGQPAICIKOE 225
DB 352 -----NAGVGSANQTEYSHHCRSYSDPNKVCCEKPKCKHGAETOSPVANLMQCKYE 406
QY 226 GPDCACLURCPVCYKIEVVNTGSAIARNVTVDNPVDGYSHASQRVLSFNLGDMRPGDKK 285
DB 407 --DCA-----GDLK--SDKK 417
QY 286 VFTVEFCPQRRGQITNV-----ATVTYCGGHKCSANVTIVNEP--CVQVNI SGADWSYV 338
DB 418 RLEL--WCDEK-----VNVCCKDIGSTVTKDKHGTCLDYATPLYNEPKCDV--VEGESGV 471
QY 339 CKPVEYSISVSNPGDLVLHDVVQIDTLP SGVTVLEAPGGEICCNKVVVRIKEMCPGETL- 397
DB 472 CK-----TKGG-----VCREGHCCPSLTLT 491
QY 398 -----QFKLVVKAQVPGRF-----TNOVAVTSESNGTC 426
DB 492 IAPSGNGTESATPTLGPYPYLTNYPDCDANKPIPSQFSTYAFCDPDTNRVGI LGRHLL--- 548
QY 427 TSCAETTHWKGLAATHM--CVLDTNDPICVGENTVYRIC 464
DB 549 --TGERTEVKGSGACSSNNKCKSGT---VCVYVNI NKHVC 583

Search completed: May 25, 2002, 22:16:36
Job time: 5270 sec

p29579 methanobact
Q28295 canis famil
P98092 bombyx mori
O14246 homo sapien
Q9tv36 sus scrofa
P34576 caenorhabdi
Q07008 rattus norv
Q05793 mus musculu
P98160 homo sapien
Q02817 homo sapien
O62918 rattus norv
P39712 saccharomyc

34 104.5 3.5 499 1 YP25_METTE
35 104.5 3.5 2813 1 VWF_CANFA
36 104.5 3.5 3133 1 HMCT_BOMMO
37 104 3.5 886 1 EMRI_HUMAN
38 104 3.5 2871 1 FBNI_PIG
39 104 3.5 3051 1 YN33_CAEEL
40 103 3.5 2531 1 NTC1_RAT
41 102.5 3.5 3707 1 PGBM_MOUSE
42 102.5 3.5 4393 1 PGBM_HUMAN
43 102.5 3.5 5179 1 MUC2_HUMAN
44 102 3.5 816 1 NEL2_RAT
45 102 3.5 1322 1 YAG3_YEAST

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

DM protein - protein search, using sw model

Run on: May 25, 2002, 22:16:42 ; Search time 39.38 Seconds
(without alignments)
546.675 Million cell updates/sec

Title: US-09-523-647-2

Perfect score: 2951

Sequence: 1 MSKLIRRVVTVALTSMASC.....ILSDTLTSPVSDTENTHYV 556

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2951	100.0	556	1 OM6_CHLPN	P23700 chlamydia p
2	2544.5	86.2	557	1 OM6_CHLPS	P23701 chlamydia p
3	2162.5	73.3	547	1 OM6C_CHLTR	P23603 chlamydia t
4	2161.5	73.2	547	1 OM6C_CHLTR	P26758 chlamydia t
5	2160.5	73.2	547	1 OM6D_CHLTR	P18151 chlamydia t
6	2129.5	72.2	547	1 OM6L_CHLTR	P21354 chlamydia t
7	130.5	4.4	1408	1 SERR_DROME	P18168 drosophila
8	129	4.4	772	1 ANNU_SCHAM	P52183 schistocerc
9	123.5	4.2	2139	1 GRB_DROME	P10040 drosophila
10	123	4.2	1295	1 GLP1_CAEEL	P13508 caenorhabdi
11	115.5	3.9	577	1 TRBM_MOUSE	P15306 mus musculu
12	115.5	3.9	2911	1 FBNI_HUMAN	P35556 homo sapien
13	113.5	3.8	555	1 DP87_DICDI	Q04503 dictyosteli
14	113.5	3.8	1808	1 TENA_CHICK	P10039 gallus gall
15	112	3.8	2907	1 FBNI_MOUSE	Q61555 mus musculu
16	112	3.8	2871	1 FBNI_MOUSE	Q29116 sus scrofa
17	111.5	3.8	1746	1 TENA_PIG	P24821 homo sapien
18	111	3.8	2201	1 FBNI_HUMAN	P35555 homo sapien
19	110.5	3.7	2871	1 FBNI_HUMAN	Q03610 caenorhabdi
20	110	3.7	1416	1 YN81_CAEEL	P07207 drosophila
21	110	3.7	2703	1 NOTC_DROME	P98133 bos taurus
22	109.5	3.7	2871	1 FBNI_BOVIN	Q99088 xenopus lae
23	109	3.7	892	1 LDL2_XENLA	P17053 paramecium
24	108	3.7	2704	1 GL68_PARP	Q9uh18 homo sapien
25	107	3.6	967	1 ATSL_HUMAN	Q03376 chironomus
26	107	3.6	1700	1 BAR3_CHITE	P98165 gallus gall
27	106.5	3.6	863	1 LDVR_CHICK	P03517 punta toro
28	106.5	3.6	1313	1 VGLM_PTPV	P03517 punta toro
29	106.5	3.6	4655	1 LRP2_HUMAN	P98164 homo sapien
30	106	3.6	1246	1 YMV2_CAEEL	P34504 caenorhabdi
31	105.5	3.6	1680	1 FUR2_DROME	P30432 drosophila
32	105	3.6	967	1 NTC1_RAT	Q9uwl1 rattus norv
33	105	3.6	2444	1 ATSL_HUMAN	P46531 homo sapien

ALIGNMENTS

RESULT 1

ID	OM6_CHLPN	STANDARD;	PRT;	556 AA.
AC	P23700; 09JO16;			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-NOV-1991 (Rel. 20, Last sequence update)			
DE	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	60 kDa outer membrane protein precursor (Cysteine-rich outer membrane protein) (CRP) (60 kDa cysteine-rich OMP).			
DE	OMCB OR OMP2 OR CPN0557 OR CP0195.			
GN	Chlamydia pneumoniae (Chlamydia pneumoniae).			
OS	Chlamydia pneumoniae (Chlamydia pneumoniae).			
OC	Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.			
OX	NCBI_TaxID=83558;			
[1]	SEQUENCE FROM N.A.			
RP	STRAIN=IOL-207;			
RC	MEDLINE=90384850; PubMed=2402463;			
RA	Watson M.W., Al-Nahdawi S., Lamden P.R., Clarke I.N.;			
RT	"The nucleotide sequence of the 60 kDa cysteine rich outer membrane protein of Chlamydia pneumoniae strain IOL-207.";			
RL	Nucleic Acids Res. 18:5299-5299(1990).			
[2]	SEQUENCE FROM N.A.			
RP	STRAIN=CWL029;			
RC	MEDLINE=99206606; PubMed=10192388;			
RA	Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,			
RT	Olinger L., Grimwood J., Davis R.W., Stephens R.S.;			
RL	"Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";			
[3]	Nat. Genet. 21:385-389(1999).			
[3]	SEQUENCE FROM N.A.			
RP	STRAIN=AR39;			
RC	MEDLINE=20350255; PubMed=10684935;			
RA	Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., Bass S.,			
RT	White O., Hickey E.K., Peterson J., Utterback T., Berry K., Dodson R.,			
RA	Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,			
RA	Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,			
RT	Eisen J., Fraser C.M.;			
RL	"Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AK39.";			
RL	Nucleic Acids Res. 28:1397-1406(2000).			
[4]	SEQUENCE FROM N.A.			
RP	STRAIN=J138;			
RC	MEDLINE=20330349; PubMed=10871362;			
RA	Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,			
RT	Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;			
RA	"Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CWL029 from USA.";			
RL	Nucleic Acids Res. 28:2311-2314(2000).			
[5]	SEQUENCE FROM N.A.			
RP	STRAIN=J138;			
RC	Shirai M., Hirakawa H., Ouchi K., Tabuchi M., Kishi F., Kimoto M.,			
RT	Takeuchi A., Nishida J., Shibata K., Fujinaga R., Yoneda H.,			
RA	Matsushima H., Tanaka C., Furukawa S., Miura K., Nakazawa A.,			

RA Ishii K., Shiba T., Hattori M., Kuhara S.;
 RT "Comparison of outer membrane protein genes omp and pmp in the whole
 RT US.";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES
 CC (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL
 CC INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT
 CC VIRULENCE FACTOR.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X53511; CAA37590.1; -;
 CC EMBL; AE001640; AAD18697.1; -;
 CC EMBL; AE002180; AAF38068.1; -;
 CC EMBL; AP002547; BAA98763.1; -;
 CC EMBL; AB033786; BAA85939.1; -;
 CC PIR; S12602; S12602.
 CC PHCI-2DPAGE; P23700; -;
 CC TIGR; CP0195; -;
 CC InterPro; IPR003506; Chlam_OMP6.
 CC PRINTS; PR01336; CHLAMIDIAOM6.
 CC Outer membrane; Transmembrane; Signal; Virulence; Complete proteome.
 KW SIGNAL 1 22 POTENTIAL.
 FT PROPEP 23 40 POTENTIAL.
 FT CHAIN 41 556 60 KDA OUTER MEMBRANE PROTEIN.
 SQ SEQUENCE 556 AA; 59719 MW; 8D7ED9234CC99458 CRC64;

Query Match 100.0%; Score 2951; DB 1; Length 556;
 Best Local Similarity 100.0%; Pred. No. 1.5e-213;
 Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSKLIRVVTVLALTSWASCFASGIEAAVAESLITKIVASAEKTPAPVPMATAKRVLR 60
 Db 1 MSKLIRVVTVLALTSWASCFASGIEAAVAESLITKIVASAEKTPAPVPMATAKRVLR 60
 QY 61 RNKQPVQKSRGAFCDKEFFVCEGRQCPVEAQOESCGRLYSVKVNDCNVEICQSVPE 120
 Db 61 RNKQPVQKSRGAFCDKEFFVCEGRQCPVEAQOESCGRLYSVKVNDCNVEICQSVPE 120
 QY 121 YATVGSPIPTIEILAIGKDCVDVITQOLPCEAEFVSSDPETPTSDGKLVWKIDRLGAG 180
 Db 121 YATVGSPIPTIEILAIGKDCVDVITQOLPCEAEFVSSDPETPTSDGKLVWKIDRLGAG 180
 QY 181 DKCKITVWKPLKGGCCFTATVACAPCLRSYKCGQPAICIKOEGDPCACLRCPVCYKI 240
 Db 181 DKCKITVWKPLKGGCCFTATVACAPCLRSYKCGQPAICIKOEGDPCACLRCPVCYKI 240
 QY 241 EVVNTGSATARNVTVDNVPDGVSHASQGVLSFNLGDMRPGDKKFTVEFCPORRQGIT 300
 Db 241 EVVNTGSATARNVTVDNVPDGVSHASQGVLSFNLGDMRPGDKKFTVEFCPORRQGIT 300
 QY 301 NVATVTCGHHKCSANVTTVNPEPCVQVNTSGADWSVCKPVEYSISVSNPGDLVLHDV 360
 Db 301 NVATVTCGHHKCSANVTTVNPEPCVQVNTSGADWSVCKPVEYSISVSNPGDLVLHDV 360
 QY 361 IQDTLPVSGVTLEAPGGEICCNKVVRIKEMCPGETLQFKLVVKAQVPGRTNOVATSE 420
 Db 361 IQDTLPVSGVTLEAPGGEICCNKVVRIKEMCPGETLQFKLVVKAQVPGRTNOVATSE 420
 QY 421 SNGCTCTCAETTHWKGLAATHMCLVLDNDPFCVGENTVYRICVTVNRGSAEDTNSLIL 480
 Db 421 SNGCTCTCAETTHWKGLAATHMCLVLDNDPFCVGENTVYRICVTVNRGSAEDTNSLIL 480
 QY 481 KFSKELOPIASSGPTKGTISGNTVVFDPALPKLGSKESVEFSVTLKGIAPGDARGEAILSS 540

Db 481 KFSKELOPIASSGPTKGTISGNTVVFDPALPKLGSKESVEFSVTLKGIAPGDARGEAILSS 540
 QY 541 DTLTSPVSDTENTHYV 556
 Db 541 DTLTSPVSDTENTHYV 556
 RESULT 2
 OM6_CHLPS STANDARD; PRT; 557 AA.
 AC P23701;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 60 kDa outer membrane protein precursor (Cysteine-rich outer membrane
 DE protein) (CRP) (60 kDa cysteine-rich OMP).
 GN OMCB OR OMP2 OR ENVB.
 OS Chlamydia psittaci (Chlamydia psittaci).
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia phila.
 OX NCBI_TaxID=83554;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EAE/A22/M;
 RX MEDLINE=90384851; PubMed=2402464;
 RA Watson M.W.; Lambden P.R.; Clarke I.N.;
 RT "The nucleotide sequence of the 60 kDa cysteine rich outer membrane
 RT protein of Chlamydia psittaci strain EAE/A22/M.";
 RL Nucleic Acids Res. 18:5300-5300(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=6BC;
 RX MEDLINE=91267949; PubMed=2050637;
 RA Everett K.D.E.; Hatch T.P.;
 RT "Sequence analysis and lipid modification of the cysteine-rich
 RT envelope proteins of Chlamydia psittaci 6BC.";
 RL J. Bacteriol. 173:3821-3830(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Watson M.W.;
 RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES
 CC (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL
 CC INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT
 CC VIRULENCE FACTOR.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X53512; CAA37592.1; -;
 CC EMBL; M61116; AAB61619.1; -;
 CC PIR; S12603; S12603.
 CC InterPro; IPR003506; Chlam_OMP6.
 CC PRINTS; PR01336; CHLAMIDIAOM6.
 CC Outer membrane; Transmembrane; Signal; Virulence.
 FT SIGNAL 1 22 POTENTIAL.
 FT PROPEP 23 40 POTENTIAL.
 FT CHAIN 41 557 60 KDA OUTER MEMBRANE PROTEIN.
 FT VARIANT 45 45 A -> S (IN STRAIN 6BC).
 FT VARIANT 73 73 E -> G (IN STRAIN 6BC).
 SQ SEQUENCE 557 AA; 59843 MW; 0D444F09EAA073C6 CRC64;

Query Match 86.2%; Score 2544.5; DB 1; Length 557;
 Best Local Similarity 85.1%; Pred. No. 4e-183;
 Matches 474; Conservative 30; Mismatches 52; Indels 1; Gaps 1;

Tue May 28 08:55:20 2002

QY 1 MSLKLRVVTVLALTSMAFCASGIERAAVAESLTKIVASAEKTPAPV-PMTAKKVLV 59
DB 1 MSLKLRVVTVLALTSMAFCASGIERAAVAESLTKIVASAEKTPAPV-PMTAKKVLV 60
QY 60 RNNKQPVKSRGAFCDKEFYCEBGRCPQVPAQOESYGRILSVKVNDDCNVEICQSV 119
DB 61 RNNKQPVKSRGAFCDKEFYCEBGRCPQVPAQOESYGRILSVKVNDDCNVEICQSV 120
QY 120 EYATVGSPIETLAIKGRKDCVDVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGA 179
DB 121 EYATVGSPIETLAIKGRKDCVDVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGA 180
QY 180 GSKCKITVWVKPLKEGCCFTAAATVACAPELRSYTKCGPAICIKQEGPCACLRCPVCYK 239
DB 181 GSKCKITVWVKPLKEGCCFTAAATVACAPELRSYTKCGPAICIKQEGPCACLRCPVCYK 240
QY 240 IEVNTGSAIARNVTVDPDGYSHASQGVLSFNLGDMRPGDKKVFVTEFCPQRGOI 299
DB 241 IEVNTGSAIARNVTVDPDGYSHASQGVLSFNLGDMRPGDKKVFVTEFCPQRGOI 300
QY 300 TNVATVYCGGHKCSANVTTVNPEPCVQVNIAGADWSYVCKPVEYSISVSNPGDLVLDV 359
DB 301 TNVATVYCGGHKCSANVTTVNPEPCVQVNIAGADWSYVCKPVEYSISVSNPGDLVLDV 360
QY 360 VIQDTLPSGVTVLEAPGGEICCNKVRIRKEMCPGETLQFKLVVKAQVPGFTNOAVTS 419
DB 361 VIQDTLPSGVTVLEAPGGEICCNKVRIRKEMCPGETLQFKLVVKAQVPGFTNOAVTS 420
QY 420 ESNCGTCSAEATTHHKGLAATHMCLVLDNDPVCVNTYRVCVNRGSAEDTNVSLI 479
DB 421 NSDCGTCSAEATTHHKGLAATHMCLVLDNDPVCVNTYRVCVNRGSAEDTNVSLI 480
QY 480 LKFSKELQPIASSGPTKGTISGNTVDFDALPKLGSKEVSFVTLKGIAPGDARGEAILS 539
DB 481 LKFSKELQPIASSGPTKGTISGNTVDFDALPKLGSKEVSFVTLKGIAPGDARGEAILS 540
QY 540 SDTLTSPVSDTENTHY 556
DB 541 SDTLTSPVSDTENTHY 557

RESULT 3
ID OM6E_CHLTR STANDARD; PRT; 547 AA.
AC P23603;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 60 kDa outer membrane protein, serovar E precursor (Cysteine-rich outer membrane protein) (60-kDa CRP).
GN OMP2 OR OMP2B.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DK 20 / SEROVAR E;
RX MEDLINE=91067486; PubMed=2251143;
RA Coles A.M., Allan I., Pearce J.H.;
RT "The nucleotide and derived amino acid sequence of the omp2 gene of Chlamydia trachomatis serovar E.";
RL Nucleic Acids Res. 18:6713-6713(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BOUR / SEROVAR E;
RX MEDLINE=91147205; PubMed=1997423;
RA de la Maza L.M., Fiedler T.J., Carlson E.J., Markoff B.A., Peterson E.M.;
RT "Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodalton protein between the trachoma and lymphogranuloma venereum biovars of Chlamydia trachomatis.";
RL Infect. Immun. 59:1196-1201(1991).
RN [3]

RP SEQUENCE FROM N.A.
RA Zhang Y.X., Caldwell H.D.;
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
CC FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT VIRULENCE FACTOR.
CC SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF STRAIN DK 20.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC EMBL: X55903; CAA39396.1; -;
CC EMBL: X54389; CAA38259.1; -;
CC EMBL: M85196; AAA23154.1; -;
CC PIR: S13120; S13120. Chlam_OMP6.
CC InterPro: IPR003506; Chlam_OMP6.
CC PRINTS: PR01336; CHLAMIDIAOM6.
CC Outer membrane; Transmembrane; Signal; Virulence.
CC SIGNAL 1 22 POTENTIAL.
CC PROPEP 23 40
CC CHAIN 41 547 60 KDA OUTER MEMBRANE PROTEIN, SEROVAR E.
CC VARIANT 33 34 SL -> FT (IN STRAIN DK20).
CC VARIANT 121 121 I -> L (IN STRAIN DK20).
CC VARIANT 132 132 V -> A (IN STRAIN DK20).
CC VARIANT 458 458 N -> S (IN STRAIN DK20).
CC SEQUENCE 547 AA; 58708 MW; 052066084F4E20AB CRC64;

Query Match 73.3%; Score 2162.5; DB 1; Length 547;
Best Local Similarity 72.4%; Pred. No. 1.5e-154;
Matches 403; Conservative 72; Mismatches 71; Indels 11; Gaps 4;
QY 1 MSLKLRVVTVLALTSMAFCASGIERAAVAESLTKIVASAEKTPAPV-PMTAKKVLV 60
DB 1 MSLKLRVVTVLALTSMAFCASGIERAAVAESLTKIVASAEKTPAPV-PMTAKKVLV 58
QY 61 RNNKQPVKSRGAFCD-KEFYCEBGRCPQVPAQOESYGRILSVKVNDDCNVEICQSV 119
DB 59 KN-----HSKETLVDKKEVAPVHESKA--TGPKQDSCFGRMYTKVNDNRNVEITQAVP 110
QY 120 EYATVGSPIETLAIKGRKDCVDVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGA 179
DB 111 EYATVGSPIETLAIKGRKDCVDVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGA 170
QY 180 GSKCKITVWVKPLKEGCCFTAAATVACAPELRSYTKCGPAICIKQEGPCACLRCPVCYK 239
DB 171 GEKSKITVWVKPLKEGCCFTAAATVACAPELRSYTKCGPAICIKQEGPCACLRCPVCYK 230
QY 240 IEVNTGSAIARNVTVDPDGYSHASQGVLSFNLGDMRPGDKKVFVTEFCPQRGOI 299
DB 231 INVNTGSAIARNVTVDPDGYSHASQGVLSFNLGDMRPGDKKVFVTEFCPQRGOI 290
QY 300 TNVATVYCGGHKCSANVTTVNPEPCVQVNIAGADWSYVCKPVEYSISVSNPGDLVLDV 359
DB 291 TNVATVYCGGHKCSANVTTVNPEPCVQVNIAGADWSYVCKPVEYSISVSNPGDLVLDV 350
QY 360 VIQDTLPSGVTVLEAPGGEICCNKVRIRKEMCPGETLQFKLVVKAQVPGFTNOAVTS 419
DB 351 VIEDTLSPGVTVLEAPGGEICCNKVRIRKEMCPGETLQFKLVVKAQVPGFTNOAVTS 410
QY 420 ESNCGTCSAEATTHHKGLAATHMCLVLDNDPVCVNTYRVCVNRGSAEDTNVSLI 479
DB 411 NSDCGTCSAEATTHHKGLAATHMCLVLDNDPVCVNTYRVCVNRGSAEDTNVSLI 470
QY 480 LKFSKELQPIASSGPTKGTISGNTVDFDALPKLGSKEVSFVTLKGIAPGDARGEAILS 539
DB 481 LKFSKELQPIASSGPTKGTISGNTVDFDALPKLGSKEVSFVTLKGIAPGDARGEAILS 539

```
Db 471 LKFSKELQPVSGPTKGTITGNTVTVFDSLPLRGSKETVEFSVTLKAVSAGDARGEALIS 530
QY 540 SDTLTPSPVSDTENTHYV 556
Db 531 SDTLTPVPSDENTHIY 547

RESULT 4
OM6C_CHLTR STANDARD; PRT; 547 AA.
AC P26758;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 60 kDa outer membrane protein, serovar C precursor (Cysteine-rich
outer membrane protein) (60-kDa CRP).
GN OMP2 OR OMP2B.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TW-3 / SEROVAR C;
RX MEDLINE=91147205; PubMed=1997423;
RA de la Maza L.M., Fiedler T.J., Carlson E.J., Markoff B.A.,
RA Peterson E.M.;
RT "Sequence diversity of the 60-kilodalton protein and of a putative
RT 15-kilodalton protein between the trachoma and lymphogranuloma
RT venereum biovars of Chlamydia trachomatis.";
RL Infect. Immun. 59:1196-1201(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Zhang Y.X., Caldwell H.D.;
RA Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES
CC (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL
CC INTEGRITY OF THE EBS OUTER ENVELOPE. IT MAY ALSO BE AN IMPORTANT
CC VIRULENCE FACTOR.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC -----
CC EMBL; X54388; CAA38257.1; -
CC DR EMBL; M85197; AAA23159.1; -
CC DR PIR; C43584; C43584.
CC DR InterPro; IPR003506; Chlam_OMP6.
CC DR PRINTS; PR01336; CHLAMIDIAOM6.
CC KW Outer membrane; Transmembrane; Signal; Virulence.
CC FT SIGNAL 1 22
CC FT PROPEP 23 40
CC FT CHAIN 41 547 60 KDA OUTER MEMBRANE PROTEIN, SEROVAR C.
CC SEQUENCE 547 AA; 58680 MW; 817BA5DC7FEA65D4 CRC64;

Query Match
Best Local Similarity 73.2%; Score 2161.5; DB 1; Length 547;
Matches 402; Conservative 74; Mismatches 68; Indels 15; Gaps 4;

QY 1 MSLKLRVVTVLALTSMAISCFASGIEAAVSAESLTKIVASAETKPAVPMTAKKVLVR 60
Db 1 MNKLIRRAVTTFATVSVASLFSAGVLETSMAESLSTNVISLADTKAK--DNTSHKSKAR 58
QY 61 RN---KQPEVKSRGAFCDKEFYPCCEGRCPQVEAQSCGRLYSVKVNDCNVVEIQCS 117
Db 59 KNHSEKTPVDR-----KEVAPVHESKA--TGPKQDSCEGRMYTVKVNDRNVEITQA 108
QY 118 VPEYATVGSYPPIELAILGKKKDCVDVVIITQQLPCEAEFVRSDDPATPTTADGKLVWKIDRL 177
```

```
Db 109 VPEYATVGSYPPIELAITATGKRCVDVVIITQQLPCEAEFVRSDDPATPTTADGKLVWKIDRL 168
QY 178 GAGDKCKITVWKPLKEGCGCFTAAATVCACPELRSYTKCGOPATCIKQEGPDCACLRCPVC 237
Db 169 GQGEKSKITVWKPLKEGCGCFTAAATVCACPEIRSVTKCGOPALCVKQEGPENACLRCPVV 228
QY 238 YKIEVNTGSAIARNVTVDNPNVDPGYSHASGORVLSFNLDGMRPDKGVFTVEFCPORRG 297
Db 229 YKINVVNOGTATARNVVENPVPDGYAHSSGQVRVLTFTGLDQPGGEBHRITITVEFCPLKRG 288
QY 298 QITNATVYTCGGHKCSANVTTVNPNPCVOVNIISGADWSYCKPVPYSISVSNPGDLVLH 357
Db 289 RATNIATVSYCGGHKNKTASVTTVINEPCVOVSIAGADWSYCKPVEYVLSVSNPGDLVLH 348
QY 358 DVVIQDTLPSGVTVLEAPGGEICCNKVVWRIKEMCPGETLQPKLVVKAQVPGRTNOVAV 417
Db 349 DVVVEDTLPSPGVTVLEAAGAQISCNKVVWTVKELNPGESLOYKVLVRAOTPGQFTNNVVV 408
QY 418 TSESNCGTCTSCAETTTTHWKGIAATHMCVLDNDPFCVGTNTVYRICVTNRGSAEDTNVS 477
Db 409 KSCSDCGTCTSCAETTVYKGVAAATHMCVVDVTDPCVCGENTVYRICVTNRGSAEDTNVS 468
QY 478 LILKFSKELQPIASSGPTKGTISGNTVTVFDPALPLGSKESVEFSVTLKGIAPGDARGEAI 537
Db 469 LMLKFSKELQPVPSGPTKGTITGNTVTVFDSLPLRGSKETVEFSVTLKAVSAGDARGEAI 528
QY 538 LSSDTLTPSPVSDTENTHYV 556
Db 529 LSSDTLTPVPSDENTHIY 547

RESULT 5
OM6D_CHLTR STANDARD; PRT; 547 AA.
AC P18151;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 60 kDa outer membrane protein precursor (Cysteine-rich outer membrane
protein) (CRP) (60 kDa cysteine-rich OMP).
GN OMCB OR OMP2 OR OMP2B OR CT443.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B/JALI20/OT;
RX MEDLINE=90128208; PubMed=2612891;
RA Watson M.W., Lambden P.R., Ward M.E., Clarke I.N.;
RT "Chlamydia trachomatis 60 kDa cysteine rich outer membrane protein:
RL sequence homology between trachoma and LGV biovars.";
RN [2]
RP FEMS Microbiol. Lett. 53:293-297(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=B/TW-05/OT;
RX MEDLINE=91141306; PubMed=2287277;
RA Allen J.E., Cerrone M.C., Beatty P.R., Stephens R.S.;
RT "Cysteine-rich outer membrane proteins of Chlamydia trachomatis
RL display compensatory sequence changes between biovariants.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/CX;
RX MEDLINE=9900809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RL Chlamydia trachomatis.";
RL Science 282:754-759(1998).
CC -1- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES
CC (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL
```


Db	679	SGSLCEEAKENCTPSPCLBQHCLNTPGYYCHCPDPRAGKHCEQURPLCSQPPCNEGCEA	738
QY	222	-----IKQEGPDC-----ACLRCPV-----CYKIEVYVNTGSAIA	250
Db	739	NVSLATSATTTTTTTTATTRKMAKPSGLPCSGHGSCEMSDVGTEFC-KCHVGHTGTGTC	797
QY	251	RNYTVDNVPV-----DG-----YSHASGQRYLS-----FNLGDMRPG--	282
Db	798	HNLNECSPNCRNGGICLDGDGDFTECEMSGWTGKRCSEATCYAGOCQNGTCTHPGAP	857
QY	283	DKKV-----FTVEFCPQRRGQTNVATVYCGGHKCSANVTVWNEPCVQVNI	333
Db	858	DKALQPHCRCAPGWTGLFCAE-----AIDQCRQCPCHNGGT-----CE-----SGA	898
QY	334	DW-SYVCKPVEYTSVSNPGLVLHDVVIODTLPSTVTVLEAPGGEICCNVVMRIKEMC	392
QY	899	GWERCVC-----AQGSGPDCRINVNCSPQPCGGGATCIDGGYGC-----IC	943
Db	944	P-----PGRHLRCELLISDPKSAQONASNTISPTALNRSQNMLDIALTG	989
QY	444	MCVLDTNPICGVGVYRICVTNRGSAEDTNVSLILKFSKLOPIASSGPTKGTISGNT	503
Db	990	RTEDDENCNACVCN-----GTSRCTNMLCGLPNKYKVDPLSKSSNLGVCQKQHE	1039
QY	504	VVFDAL 509	
Db	1040	VCVPAL 1045	

RESULT 8

ANNU_SCHAM

ID

ANNU_SCHAM

STANDARD;

PRT;

772 AA.

P52183;

01-OCT-1996 (Rel. 34, Created)

01-OCT-1996 (Rel. 34, Last annotation update)

15-JUL-1999 (Rel. 38, Last annotation update)

Annulin (protein-glutamine gamma-glutamyltransferase) (EC 2.3.2.13)

DE

Transglutaminase).

OS

Schistocerca americana (American grasshopper).

OC

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC

Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;

OC

Acridomorpha; Acridoidea; Acrididae; Schistocerca.

OX

NCBI_TaxID=7009;

[1]

SEQUENCE FROM N.A.

RP

MEDLINE=93050772; PubMed=1358727;

RX

Singer M.A., Hortsch M., Goodman C.S., Bentley D.;

RA

'Annulin, a protein expressed at limb segment boundaries in the

RT

grasshopper embryo, is homologous to protein cross-linking

RT

transglutaminases.';

RL

Dev. Biol. 154:143-159(1992).

CC

-1- FUNCTION: PARTICIPATES IN MORPHOGENETIC ACTIVITIES OF THE CELLS,

CC

MAYBE BY STABILIZING THE MEMBRANE OR SUBCORTICAL STRUCTURES OF

CC

CELLS THAT ARE UNDER MECHANICAL STRESS. PROBABLY CATALYZES THE

CC

CROSS-LINKING OF PROTEINS AND THE CONJUGATION OF POLYAMINES TO

CC

PROTEINS.

CC

-1- CATALYTIC ACTIVITY: Protein glutamine + alkylamine = protein N5-

CC

alkylglutamine + NH(3).

CC

-1- COFACTOR: CALCIUM IS REQUIRED FOR TRANSGLUTAMINASE ACTIVITY (BY

CC

SIMILARITY).

CC

-1- SUBCELLULAR LOCATION: INTRACELLULAR AND PERIPHERALLY ASSOCIATED

CC

WITH THE INNER LEAFLET OF THE CELL MEMBRANE, USING A FATTY ACID

CC

LINKAGE.

CC

-1- TISSUE SPECIFICITY: HAS AN ANNULAR, OR RING-LIKE EXPRESSION

CC

PATTERN IN EPITHELIAL ANNULI OF DEVELOPING LIMB SEGMENT BOUNDARY

CC

CELLS. IN EMBRYOS, IT IS SEEN IN GASTRULATING CELLS, IN CELLS

CC

SURROUNDING RAPIDLY DIVIDING NEUROBLASTS, AND IN MUSCLE PIONEER

CC

CELLS INVAGINATING TO FORM APODEMES.

CC

-1- DEVELOPMENTAL STAGE: EXPRESSION OF THIS PROTEIN IN EMBRYOS AND

CC

LIMBS IS ASSOCIATED WITH AREAS UNDERGOING MOVEMENTS, MORPHOGENETIC

CC

REARRANGEMENTS, OR RAPID CELL DIVISION. EXPRESSION OF ANNULIN

PRECEDES THE FIRST MORPHOLOGICAL SIGNS OF SEGMENTATION IN THE DEVELOPING LIMBS.
-!- SIMILARITY: BELONGS TO THE TRANSGLUTAMINASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: M92291; AA29806.1; -;
HSSP: P00488; IGGU.
InterPro: IPR002931; Transglut_core.
InterPro: IPR001102; Transglut_core.
Pfam: PF00927; Transglutamin_C; 1.
Pfam: PF00868; Transglutamin_N; 1.
Pfam: PF01841; Transglutamin_N; 1.
SMART: SM00460; TGC; 1.
PROSITE: PS00547; TRANSGLUTAMINASES; 1.
Transferase: Acyltransferase; Calcium-binding.
LIPID 4 4
LIPID 5 5
ACT_SITE 341 341 BY SIMILARITY.
DOMAIN 567 570 POLY-VAL.
SEQUENCE 772 AA; 85941 MW; FA5A3CE6A7C4E394 CRC64;

Query Match 4.4%; Score 129; DB 1; Length 772;
Best Local Similarity 23.7%; Pred. No. 0.036;
Matches 79; Conservative 55; Mismatches 118; Indels 82; Gaps 19;
QY 41 SAETKPAVPMTAKKVLVRNK-----QVQKSRGAFCDKEF-----YP--CEGRQCP 89
Db 476 SGTPQ-----LKLIRKMDLIGQNTSTKAVGRFQREDITNTYKYPEKSVVEERAAM 526
QY 90 VEA--QGESYGRLYSVKVNDCN-VEIQSVPEYATVGSYPPIELIAGKKDCVDVVIT 146
Db 527 LKALRQSESLFSRY--LNEEDFNDFHNFELRDDIIVGSFVSVVVYKNSNOODTVT 583
QY 147 QQLPCEAEFVSSDPETPTSDGLVWKIDRL---GAGDKKITV---WVKPLKEGCGFT 199
Db 584 VLLRVDIVLY-----TGHVKGDKGKKEVERLIRAGAVEEVRIDVSYEDYKHLVDQCAFN 638
QY 200 AATVCAPELRSYTKCGQ-----PAICIKQGPDCACLRCPVYKIEVVNTGSAARN 252
Db 639 IACLATVHD-TNVEYFAQDDFRVRKPKDIKLEGE-----PV-----QGQMSAV 682
QY 253 VTVDNVPV-----DCYSHASQGV-LSFNLGDMRFGDKKVFTEFCPQRRGQITN 301
Db 683 ATKLNPLPIPVKQGLIEGPGIAKTQKIKLSQIA---PGEASVNFKFTPKYDGR--- 736
QY 302 VAPVTYCGGHKCSANTVTTVN---EPCVQVNISS 332
Db 737 -ATIAKFSKSLDLDVGDGLNFVFPKKEVNGTG 769

RESULT 9
ID CRB_DROME STANDARD; PRT; 2139 AA.
AC P10040;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Crumbs protein precursor (95F).
GN CRB.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_taxid=7227;
RN [1]

SEQUENCE FROM N.A.
STRAIN-OREGON-R; TISSUE-Embryo;
MEDLINE-90263104; PubMed-2344615;
Tepass U., Theres C., Knust E.;
"Crumbs encodes an EGF-like protein expressed on apical membranes of
Drosophila epithelial cells and required for organization of
epithelia.";
Cell 61:787-799(1990).
RN [2]
RP SEQUENCE OF 1663-1955 FROM N.A.
RC TISSUE-Embryo;
RX MEDLINE-87218537; PubMed-3107986;
RA Knust E., Dietrich U., Tepass U., Brener K.A., Weigel D.,
Vaessin H., Campos-Ortega J.A.;
"EGF homologous sequences encoded in the genome of Drosophila
melanogaster, and their relation to neurogenic genes.";
EMBO J. 6:761-766(1987).
RL Cell 61:787-799(1990).
CC -!- FUNCTION: MAY PLAY A ROLE IN THE DEVELOPMENT OF EPITHELIA,
POSSIBLY FOR THE ESTABLISHMENT AND/OR MAINTENANCE OF CELL
POLARITY. IT MAY ACT AS A SIGNAL.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- PTM: PHOSPHORYLATED IN THE CYTOPLASMIC DOMAIN (POTENTIAL).
CC -!- SIMILARITY: CONTAINS 29 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 3 LAMININ G-LIKE DOMAINS.
CC -----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: M33753; AAA28428.1; ALT_SEQ.
EMBL: X05144; CAA28793.1; -;
PIR: B26637; B26637.
PIR: A35672; A35672.
HSSP: P00740; LIXA.
FlyBase: FBgn0000368; crb.
InterPro: IPR000152; Aax_hydroxyl.
InterPro: IPR000561; EGF-like.
InterPro: IPR000742; EGF_2.
InterPro: IPR001881; EGF_Ca.
InterPro: IPR001438; EGF-II.
InterPro: IPR001791; Laminin_G.
Pfam: PF00008; EGF; 26.
Pfam: PF00054; laminin_G; 3.
PRINTS: PR00010; EGFBL00D.
SMART: SM00179; EGF_CA; 11.
SMART: SM00001; EGF_like; 16.
SMART: SM00282; Lamg; 3.
PROSITE: PS00010; ASX_HYDROXYL; 15.
PROSITE: PS00022; EGF_1; 26.
PROSITE: PS01186; EGF_2; 17.
PROSITE: PS01187; EGF_CA; 15.
PROSITE: PS00025; LAM_G_DOMAIN; 3.
Differentiation: Repeat; EGF-like domain; Transmembrane;
Glycoprotein; Signal; Phosphorylation.
KW SIGNAL 1 90
FT CHAIN 91 2139 CRUMBS PROTEIN.
FT DOMAIN 91 2084 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 2085 2111 POTENTIAL.
FT DOMAIN 2112 2139 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 267 303 EGF-LIKE 1.
FT DOMAIN 306 343 EGF-LIKE 2.
FT DOMAIN 348 386 EGF-LIKE 3.
FT DOMAIN 388 425 EGF-LIKE 4.
FT DOMAIN 427 463 EGF-LIKE 5.
FT DOMAIN 464 500 EGF-LIKE 6.
FT DOMAIN 501 532 EGF-LIKE 7.
FT DOMAIN 545 581 EGF-LIKE 8.
FT DOMAIN 582 611 EGF-LIKE 9.
FT DOMAIN 609 646 EGF-LIKE 10.

FT	DOMAIN	648	685	EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	946	957	BY SIMILARITY.
FT	DOMAIN	687	723	EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	952	966	BY SIMILARITY.
FT	DOMAIN	725	761	EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	968	977	BY SIMILARITY.
FT	DOMAIN	763	800	EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	984	995	BY SIMILARITY.
FT	DOMAIN	802	838	EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	989	1009	BY SIMILARITY.
FT	DOMAIN	840	902	EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	1011	1020	BY SIMILARITY.
FT	DOMAIN	904	940	EGF-LIKE 17, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	1211	1222	BY SIMILARITY.
FT	DOMAIN	942	978	EGF-LIKE 18, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	1216	1231	BY SIMILARITY.
FT	DOMAIN	980	1021	EGF-LIKE 19, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	1233	1242	BY SIMILARITY.
FT	DOMAIN	1023	1205	EGF-LIKE 20, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	1485	1496	BY SIMILARITY.
FT	DOMAIN	1207	1243	LAMININ G-LIKE 1.	FT	DISULFID	1490	1505	BY SIMILARITY.
FT	DOMAIN	1250	1480	LAMININ G-LIKE 2.	FT	DISULFID	1507	1516	BY SIMILARITY.
FT	DOMAIN	1481	1517	LAMININ G-LIKE 3.	FT	DISULFID	1763	1774	BY SIMILARITY.
FT	DOMAIN	1558	1758	EGF-LIKE 21.	FT	DISULFID	1768	1783	BY SIMILARITY.
FT	DOMAIN	1759	1795	EGF-LIKE 22.	FT	DISULFID	1785	1794	BY SIMILARITY.
FT	DOMAIN	1797	1833	EGF-LIKE 23, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	1801	1812	BY SIMILARITY.
FT	DOMAIN	1835	1871	EGF-LIKE 24, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	1806	1821	BY SIMILARITY.
FT	DOMAIN	1874	1915	EGF-LIKE 25.	FT	DISULFID	1823	1832	BY SIMILARITY.
FT	DOMAIN	1915	1951	EGF-LIKE 26.	FT	DISULFID	1839	1850	BY SIMILARITY.
FT	DOMAIN	1953	1989	EGF-LIKE 27, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	1844	1859	BY SIMILARITY.
FT	DOMAIN	1991	2029	EGF-LIKE 28, CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	1861	1870	BY SIMILARITY.
FT	DOMAIN	2030	2070	EGF-LIKE 29.	FT	DISULFID	1878	1889	BY SIMILARITY.
FT	DISULFID	271	282	BY SIMILARITY.	FT	DISULFID	1883	1903	BY SIMILARITY.
FT	DISULFID	276	291	BY SIMILARITY.	FT	DISULFID	1905	1914	BY SIMILARITY.
FT	DISULFID	293	302	BY SIMILARITY.	FT	DISULFID	1919	1930	BY SIMILARITY.
FT	DISULFID	310	321	BY SIMILARITY.	FT	DISULFID	1924	1939	BY SIMILARITY.
FT	DISULFID	315	331	BY SIMILARITY.	FT	DISULFID	1941	1950	BY SIMILARITY.
FT	DISULFID	333	342	BY SIMILARITY.	FT	DISULFID	1957	1968	BY SIMILARITY.
FT	DISULFID	352	363	BY SIMILARITY.	FT	DISULFID	1962	1977	BY SIMILARITY.
FT	DISULFID	357	374	BY SIMILARITY.	FT	DISULFID	1979	1988	BY SIMILARITY.
FT	DISULFID	376	385	BY SIMILARITY.	FT	DISULFID	1995	2008	BY SIMILARITY.
FT	DISULFID	392	403	BY SIMILARITY.	FT	DISULFID	2002	2017	BY SIMILARITY.
FT	DISULFID	397	412	BY SIMILARITY.	FT	DISULFID	2019	2028	BY SIMILARITY.
FT	DISULFID	414	424	BY SIMILARITY.	FT	CARBOHYD	37	37	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	431	442	BY SIMILARITY.	FT	CARBOHYD	96	96	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	436	451	BY SIMILARITY.	FT	CARBOHYD	198	198	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	453	462	BY SIMILARITY.	FT	CARBOHYD	238	238	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	468	479	BY SIMILARITY.	FT	CARBOHYD	239	239	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	473	488	BY SIMILARITY.	FT	CARBOHYD	336	336	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	490	499	BY SIMILARITY.	FT	CARBOHYD	400	400	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	505	515	BY SIMILARITY.	FT	CARBOHYD	550	550	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	509	520	BY SIMILARITY.					
FT	DISULFID	532	531	BY SIMILARITY.					
FT	DISULFID	549	562	BY SIMILARITY.					
FT	DISULFID	556	569	BY SIMILARITY.					
FT	DISULFID	571	580	BY SIMILARITY.					
FT	DISULFID	586	597	BY SIMILARITY.					
FT	DISULFID	591	602	BY SIMILARITY.					
FT	DISULFID	604	610	BY SIMILARITY.					
FT	DISULFID	613	624	BY SIMILARITY.					
FT	DISULFID	618	634	BY SIMILARITY.					
FT	DISULFID	636	645	BY SIMILARITY.					
FT	DISULFID	652	664	BY SIMILARITY.					
FT	DISULFID	659	673	BY SIMILARITY.					
FT	DISULFID	675	684	BY SIMILARITY.					
FT	DISULFID	691	702	BY SIMILARITY.					
FT	DISULFID	696	711	BY SIMILARITY.					
FT	DISULFID	713	722	BY SIMILARITY.					
FT	DISULFID	729	740	BY SIMILARITY.					
FT	DISULFID	734	749	BY SIMILARITY.					
FT	DISULFID	751	760	BY SIMILARITY.					
FT	DISULFID	767	778	BY SIMILARITY.					
FT	DISULFID	772	787	BY SIMILARITY.					
FT	DISULFID	789	799	BY SIMILARITY.					
FT	DISULFID	806	817	BY SIMILARITY.					
FT	DISULFID	811	826	BY SIMILARITY.					
FT	DISULFID	828	837	BY SIMILARITY.					
FT	DISULFID	844	855	BY SIMILARITY.					
FT	DISULFID	849	890	BY SIMILARITY.					
FT	DISULFID	892	901	BY SIMILARITY.					
FT	DISULFID	908	919	BY SIMILARITY.					
FT	DISULFID	913	928	BY SIMILARITY.					
FT	DISULFID	930	939	BY SIMILARITY.					

Query Match 4.2%; Score 123.5; DB 1; Length 2139;
Best Local Similarity 20.3%; Pred. No. 0.3;
Matches 74; Conservative 34; Mismatches 120; Indels 137; Gaps 19;
QY 72 GAFCDKEFYPCER-----GRC-----OPVEAQOQESCY--GRLY 102
Db 421 GAFCDQNVDECDKRNPCNLNGRCLHTYGYTCQCLDGHGGEICDRPMTCTQOCFNGGTCL 480
QY 103 SVKVNDDC-----NVEICQSVPEYATVGSYPYPIEILAIAGKDCVDVVITQOLPCEAFV 156
Db 481 DKPIGFQCLCPPEYTGELCQIAPSCA---QQCPIDSECVGGK-CVCKPGSSGVNCQ---- 532
QY 157 SSDPETTPSDGKLVKIDRLGAGDKCKITVWVKPLKEGCGCFTAAATVCAC----- 206
Db 533 -----TSTGDGASALALPTI---NCNATNG-KCLNGGTCSMNGTHCYCAVGVSGDRCE 581
QY 207 -PELRSYTKGQPAICIKQEGPDCACLRCPVCYKIEVVNTGSALARNVTVDPVPGYSH 265
Db 582 KAENCPLNCOEPMVCVQNO---CLCPENKVC-----NQCATQPCQNG--- 621
QY 266 ASQORVLSFNLGDMRFGD-----DLPNGDYECKCTRGWTGTCGNDVDECTLHPKICNGICKNEKGYKC 673
Db 622 --GECV-----DLPNGDYECKCTRGWTGTCGNDVDECTLHPKICNGICKNEKGYKC 673
QY 302 VATVTYCGGHKCSANVTTVVNEPCVQVNIAGD-----WSYVCKP-----VEYSI 346
Db 674 YCTPGTGVH-CDSVDVDECLSFPC-----LNGATCHNKINAYECVQCPGYEGENCEVDIDE 728
QY 347 SVSNP 351
Db 729 CGSNP 733


```

RESULT 10
GLP1_CAEEL
ID GLP1_CAEEL STANDARD; PRT; 1295 AA.
AC PI3508;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Glp-1 protein precursor.
GN GLP-1 OR EMB-33 OR F02A9.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=89336787; PubMed=2758466;
RA Yochem J., Greenwald I.;
RT "Glp-1 and lin-12, genes implicated in distinct cell-cell
RL interactions in C. elegans, encode similar transmembrane proteins.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Karshaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL elegans";
RN [3]
RP Nature 368:32-38(1994).
DE DELETION OF 1174-1295.
RX MEDLINE=91351288; PubMed=1881436;
RA Mango S.E., Maine E.M., Kimble J.;
RT "Carboxy-terminal truncation activates glp-1 protein to specify
RL vulval fates in Caenorhabditis elegans.";
RN [4]
RP Nature 352:811-815(1991).
DE CHARACTERIZATION OF FUNCTION OF THE ANK-REPEATS.
RX MEDLINE=93354444; PubMed=8350921;
RA Roehl H., Kimble J.;
RT "Control of cell fate in C. elegans by a GLP-1 peptide consisting
RL primarily of ankyrin repeats.";
RN [5]
RP Nature 364:632-635(1993).
DE FUNCTION.
RX MEDLINE=94208066; PubMed=8156602;
RA Mello C.C., Draper B.W., Priess J.R.;
RT "The maternal genes apx-1 and glp-1 and establishment of
RL dorsal-ventral polarity in the early C. elegans embryo.";
RN [6]
RP Cell 77:95-106(1994).
DE FUNCTION: INVOLVED IN THE SPECIFICATION OF THE CELL FATES OF THE
CC BLASTOMERES, ABA AND APA. PROPER SIGNALING BY GLP-1 INDUCES ABA
CC DESCENDANTS TO PRODUCE ANTERIOR PHARYNGEAL CELLS, AND APA
CC ESTABLISHMENT OF THE DORSAL-VENTRAL AXIS IN EARLY EMBRYOS.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DEVELOPMENTAL STAGE: ACTS ON ABA DEVELOPMENT DURING 4-CELL AND
CC 12-CELL STAGES, AND ON ABA DEVELOPMENT DURING 12-CELL AND 28-CELL
CC STAGES.
CC -1- SIMILARITY: HIGH, TO C.ELEGANS LIN-12.
CC -1- SIMILARITY: CONTAINS 10 EGF-LIKE DOMAINS.

```

```

CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M25580; AAA28058.1; -
CC EMBL; Z19555; CAA79620.1; -
CC EMBL; Z29116; CAA79620.1; JOINED.
CC EMBL; Z29116; CAA82373.1; -
CC EMBL; Z19555; CAA82373.1; JOINED.
CC PIR; A32901; A32901.
CC HSSP; P00740; IEDM.
CC WormPep; F02A9.6; CE00237.
CC InterPro; IPR002110; ANK.
CC InterPro; IPR000152; Asx_hydroxyl.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR000800; Notch.
CC Pfam; PF00023; ank; 6.
CC Pfam; PF00008; EGF; 10.
CC Pfam; PF00066; notch; 3.
CC PRINTS; PR01452; NOTCH.
CC SMART; SM00248; ANK; 4.
CC SMART; SM00179; EGF_Ca; 1.
CC SMART; SM00001; EGF_Like; 7.
CC SMART; SM00004; NL; 3.
CC PROSITE; PS50088; ANK_REPEAT; 3.
CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
CC PROSITE; PS00010; ASX_HYDROXYL; 2.
CC PROSITE; PS00022; EGF_1; 10.
CC PROSITE; PS01186; EGF_2; 8.
CC PROSITE; PS01187; EGF_Ca; 1.
CC Differentiation; Repeat; ANK repeat; EGF-like domain; Transmembrane;
KW Glycoprotein; Signal.
FT SIGNAL 1 15
FT CHAIN 16 1295
FT DOMAIN 16 764
FT TRANSMEM 765 786
FT DOMAIN 787 1295
FT DOMAIN 19 58
FT DOMAIN 117 152
FT DOMAIN 154 190
FT DOMAIN 190 230
FT DOMAIN 232 269
FT DOMAIN 271 308
FT DOMAIN 316 359
FT DOMAIN 369 406
FT DOMAIN 407 443
FT DOMAIN 446 479
FT REPEAT 493 527
FT REPEAT 528 568
FT REPEAT 569 608
FT REPEAT 961 990
FT REPEAT 994 1023
FT REPEAT 1030 1062
FT REPEAT 1074 1103
FT REPEAT 1107 1136
FT DISULFID 23 35
FT DISULFID 29 46
FT DISULFID 48 57
FT DISULFID 121 131
FT DISULFID 126 140
FT DISULFID 142 151
FT DISULFID 158 169
FT DISULFID 163 178
FT DISULFID 180 189

```


FT	DISULFID	194	206	BY SIMILARITY.
FT	DISULFID	201	218	BY SIMILARITY.
FT	DISULFID	220	229	BY SIMILARITY.
FT	DISULFID	236	248	BY SIMILARITY.
FT	DISULFID	242	257	BY SIMILARITY.
FT	DISULFID	259	268	BY SIMILARITY.
FT	DISULFID	275	286	BY SIMILARITY.
FT	DISULFID	280	296	BY SIMILARITY.
FT	DISULFID	298	307	BY SIMILARITY.
FT	DISULFID	373	384	BY SIMILARITY.
FT	DISULFID	378	394	BY SIMILARITY.
FT	DISULFID	396	405	BY SIMILARITY.
FT	DISULFID	411	422	BY SIMILARITY.
FT	DISULFID	416	431	BY SIMILARITY.
FT	DISULFID	433	442	BY SIMILARITY.
FT	DISULFID	450	461	BY SIMILARITY.
FT	DISULFID	455	467	BY SIMILARITY.
FT	DISULFID	469	478	BY SIMILARITY.
FT	CARBOHYD	244	244	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	245	245	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	333	333	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	381	381	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	609	609	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	675	675	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	1295 AA; 144078 MW; 422AAD0A2DEEF3B4 CRC64;		
Query Match 4.2%; Score 123; DB 1; Length 1295;				
Best Local Similarity 19.7%; Pred. No. 0.19;				
Matches 108; Conservative 52; Mismatches 182; Indels 206; Gaps 34;				
QY	72	GAFCKEFP----	CEGRQPVFAQ-----	ESCGRLYSVKVNDNCVE-----113
Db	186	GRYCETCALMGNICNHCRCIPNRDEK	NFRVCDSGYEGEFCNKDKNECLIEETCVNN	245
QY	114	-----	ICQSPYATVGSPIETLAIGKDCVDVI-----	TOQLP--C 151
Db	246	STCFNLHGDFCTCKPGYA--	GKYCEALDM--CKDYVCQDGYCAHDSNQMPICYC	298
QY	152	EAEFVSSDPE-TTPTS	SDGLWKIDRLGAGDKCKITVWKPLK-----	EGCC 197
Db	299	EQGFTGQRCIECPSGFG-----	GIHCDL-----	PLQRPHCSRNGTCYNDGRC 342
QY	198	FTAAATVC-----	AC-----	PELRS--YTKGQPAIGI--KOEGRDCAAC-----231
Db	343	INGFCVCEPDYIGDRCEINRKDFPD	IQSCKYNPCVNNATCIDLKNSGYSCHCPLGYG	402
QY	232	LRCP----	VCYKIEVNTGSAIARNVTVDNVPDGYSHASQORVLSFNLGDMRPGDKKVF	288
Db	403	LNCEOHLCTPTTCANGGTCEGVNGVIR	CNPGFS-----	GD-----440
QY	289	VEFCPQRRGQITVATVYTCGGHKCSANVT	VVNEPCVQVNVISGADWSYVCKPVEYSISV	348
Db	441	--YCEIKDROL-----	CSRHPCKNGV-----	CKNTGYCECQYGTGPTCEEVLVI 484
QY	349	SNPGDLVLDHWIQ----	DPLPSGVTVLEAPGGBICCNVYVRIKEM-----	CPGETLQ 398
Db	485	EKSKEVIRDLCEQRKCMDLASNGICNPE-----	CN-----	LEECNFDGDCSGGQRP 532
QY	399	FKLVVKAQVGRFTNQAVTS	ESNCGTCT-SCAETTHWKL-----	AATHM---CV 446
Db	533	FS-----	KQYPARCADQFA-----	NGVCNOECNNEECLDGLDQCSELFRCPAHRIKCI 583
QY	447	LDTNDPTC-----	VGENTVYRICVTNRGSAEDTNVSLIL---	KFSKELQPI---ASSGP 494
Db	584	ERRDGVNLECSFIG-----	CGFDGDCNNGTEAILLSDIRIKVQIDPIERQATGGE	636
QY	495	TKGTISGN 502		
Db	637	TLMQISAN 644		
RESULT 11				

TRBM_MOUSE	TRBM_MOUSE	STANDARD;	PRT;	577 AA.
ID	PL5306;			
AC	01-APR-1990 (Rel. 14, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Thrombomodulin precursor (Fetomodulin) (TM).			
GN	THBD.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RA	SEQUENCE FROM N.A.			
RP	MEDLINE=89128454; PubMed=2536925;			
RA	Dittman W.A., Majerus P.W.;			
RT	"Sequence of a cDNA for mouse thrombomodulin and comparison of the			
RT	predicted mouse and human amino acid sequences.";			
RL	Nucleic Acids Res. 17:802-802(1989).			
RN	SEQUENCE FROM N.A.			
RP	MEDLINE=89008498; PubMed=2844823;			
RA	Dittman W.A., Kumada T., Sadler J.E., Majerus P.W.;			
RT	"The structure and function of mouse thrombomodulin. Phorbol			
RT	myristate acetate stimulates degradation and synthesis of			
RT	thrombomodulin without affecting mRNA levels in hemangioma cells.";			
RL	J. Biol. Chem. 263:15815-15822(1988).			
CC	-!- FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHELIAL CELL RECEPTOR			
CC	THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS			
CC	COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE			
CC	ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA			
CC	SCISSONS THE ACTIVATED COFACTORS OF THE COAGULATION MECHANISM,			
CC	FACTOR VA AND FACTOR VIIIA, AND THEREBY REDUCES THE AMOUNT OF			
CC	THROMBIN GENERATED.			
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-!- TISSUE SPECIFICITY: ENDOTHELIAL CELLS ARE UNIQUE IN SYNTHESIZING			
CC	THROMBOMODULIN (BY SIMILARITY).			
CC	-!- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
DR	EMBL; X14432; CAA32597.1;			
DR	PIR; A32001; A32001.			
DR	PIR; S08488; S08488.			
DR	HSSP; P07204; LEGT.			
DR	MGD; MGI:98736; Thbd.			
DR	InterPro; IPR000152; Asx_hydroxyl.			
DR	InterPro; IPR000561; EGF-like.			
DR	InterPro; IPR001881; EGF_Ca.			
DR	InterPro; IPR001491; Thrombomodulin.			
DR	InterPro; IPR001304; lectin_c.			
DR	Pfam; PF00008; EGF; 4.			
DR	Pfam; PF00059; lectin_c; 1.			
DR	PRINTS; PR00907; THROMBOMODULN.			
DR	SMART; SM00034; CLECT; 1.			
DR	SMART; SM00179; EGF_CA; 1.			
DR	SMART; SM00001; EGF_like; 5.			
DR	PROSITE; PS00010; ASX_HYDROXYL; 2.			
DR	PROSITE; PS00022; EGF_1; FALSE_NEG.			
DR	PROSITE; PS01186; EGF_2; 3.			
DR	PROSITE; PS01187; EGF_CA; 2.			
DR	PROSITE; PS50041; C-TYPE-LECTIN_2; 1.			
DR	Endothelial cell; Receptor; Blood coagulation; Repeat; Transmembrane;			
KW	Glycoprotein; Signal; EGF-like domain.			
FT	SIGNAL	1	16	POTENTIAL.
FT	CHAIN	17	577	THROMBOMODULIN.
FT	DOMAIN	17	517	EXTRACELLULAR (POTENTIAL).

QY	38	IVASAETKAP-VPMTAKKVLVRRNKQVPEOKSGAFCKEY-----PCEGRCPQV	90
Db	136	VTVSTATEAAPGEAWEEK-----PCETETGFLC--EYFTASCRPLTVNTRDPE	184
QY	91	EAQOESCGRLYSKVNDNCNVEIQSVPEYATVGSXPFIETALIGKDKCDVVITQQLP	150
Db	185	AAHISSTYNTPEGVSCAD-----FQTLF--VGSSAAVEPLGL-----ELV	222
QY	151	CEAEFVSSDPEITPSPDGKLVHKIDELGADCKKITVWVKPKEGCC-----FTAAIVC	204
Db	223	CRA-----PGTSEGHAW--EATGA-----WNCSEVGGCCYLCNRSTNEPRC	264
QY	205	ACEPLRSYTKGQPAICIKQEGPDCACLRPCYKIEVVNTGSAIARNWVDN--PVPDGY	263
Db	265	LCPR-----DMDLQADGRSCA-----RPVVQSNCLCEHFCVSNAEVPGSGY	305
QY	264	S-----HASGORVLSFNLGMRPDKKVFTEVFCFQRRGQITNVTATYTCGGHKCS	314
Db	306	SCMCETGYOLAADGHR--CEDYDDCKQGPNP-----CPQ-----LCWNTKGGFECF	349
QY	315	A-NVTIVNEPCQV--NISGADWSYCKVP--EYSISVNPQ-----DLVLH	357
Db	350	CYDGYELVDGCEVLLDFCFNSCNEFCQCPSPDYR--CICAPGAPKDPDPHKCEMPCFN	408
QY	358	DYVI-QDTLPSCVTVLEAP-----GGEICCNKVVRKEMCPGETLQFKLVVRAQVPR	410
Db	409	ETSCPADCPNSPTVCECPGEFILLDEGSVCTD-----IDECQSECTSEC---RNPFGS	460
QY	411	ETNQVAVTVSESCNGCTCTCAETTHWKGLAETHMCVLDNDPDCVGENTVRIVCTNRGS	470
Db	461	Y-----ECICGPDITALAGO-----ISKDCDPIPVREDT-----KEEGS	494
QY	471	ADBTWVSLILKFSKLOQPIASSGPTKGTISG	501

```

Db      495 GE-----ppvs-----ptpgSPtG 508

RESULT 12
FBN2_HUMAN
ID      FBN2_HUMAN      STANDARD;      PRT; 2911
AC      P35556;
DT      01-JUN-1994 (Rel. 29, Created)
DI      01-FEB-1996 (Rel. 33, Last sequence update)
DE      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Fibrillin 2 precursor.
GN      FBN2.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; \
OC      Mammalia; Euthera; Primates; Catarrhini; \
OX      NCBI_TaxID=9606;
OX      [1]
RX      SEQUENCE FROM N.A.
RP      MEDLINE=94165150; PubMed=8120105;
RA      Zhang H., Apfelroth S.D., Hu W., Davis E., \
RA      Bonadio J., Mecham R.P., Ramirez F.; \
RA      "Structure and expression of fibrillin-2 \
RA      component preferentially located in elastic \
RA      J. Cell Biol. 124:855-863(1994). \
RA      [2]
RP      SEQUENCE OF 752-1505 FROM N.A.
RX      MEDLINE=91304567; PubMed=1852206;
RA      Lee B., Godfrey M., Vitaie E., Hori H., M., \
RA      Tsipouras P., Ramirez F., Hollister D.; \
RA      "Linkage of Marfan syndrome and a phenoty \
RA      two different fibrillin genes."; \
RA      Nature 352:330-334(1991). \
RA      [3]
RP      VARIANTS CCA TYR-1252 AND SER-1433, AND V \
RX      MEDLINE=96083599; PubMed=7493032;
RA      Putnam E.A., Zhang H., Ramirez F., Millev \
RA      "Fibrillin-2 (FBN2) mutations result in t \
RA      congenital contractual arachnoidactyly."; \
RA      Nat. Genet. 11:456-458(1995). \
RA      [4]
RP      VARIANTS CCA HIS-1114.
RX      MEDLINE=98407789; PubMed=9737771;
RA      Babcock D., Gasner C., Francke U., Maslen \
RA      "A single mutation that results in an esp \
RA      partial exon skipping in a family with co \
RA      arachnoidactyly."; \
RA      Hum. Genet. 103:22-28(1998). \
RA      [5]
RP      VARIANTS CCA PHE-1141 AND TRP-1252.
RX      MEDLINE=20259236; PubMed=10797416;
RA      Beilhe S., Zhou G., Wang M., Der Kalousti \
RA      Godfrey M.; \
RA      "Two novel fibrillin-2 mutations in conge \
RA      arachnoidactyly."; \
RA      Am. J. Med. Genet. 92:7-12(2000).
CC      -!- FUNCTION: STRUCTURAL COMPONENT OF CON \
CC      THAT BINDS CALCIUM. FIBRILLIN-2-CONTA \
CC      THE EARLY PROCESS OF ELASTIC FIBER AS \
CC      -!- DISEASE: DEFECTS IN FBN2 ARE THE CAUS \
CC      ARACHNOIDACTYLY (CCA) (ALSO KNOWN AS B \
CC      PHENOTYPICALLY SIMILAR TO MARFAN SYND \
CC      AORTA AND THE EYES \
CC      -!- SIMILARITY: CONTAINS 43 CALCIUM-BINDING \
CC      EGF-LIKE DOMAINS. \
CC      -!- SIMILARITY: CONTAINS 7 TGF-BETA BINDING \
CC      -----
CC      This SWISS-PROT entry is copyright. It is \
CC      between the Swiss Institute of Bioinforma \
CC      the European Bioinformatics Institute. Th \
CC      use by non-profit institutions as long \
CC      modified and this statement is not remov \
CC      entities requires a license agreement (See \
CC      or send an email to license@isb-sib.ch).

```



```
Db 317 DI---DECSIIPIGICETGECSTNTVGSFYFCVCP--RGVVTSDGSRCIDQRTGMOFSGLDV 371
Qy 230 -----ACLRCPVCVKIEWN-----TGSATARNVTVDN-----257
Db 372 GRCAQELPGRMTKMQCCCEPRGRCWIGITPEACPVRGSEYYRRLCMLDGLPMGGIPGSAGS 431
Qy 258 -----PVPDGYSHASQORVLSFNLGDMRPGDKKVF---287
Db 432 RPOGTGGNGFAPSGNGNGYGGGTGTFPIPGNGGFSFGVGAGVGAGGQGIITGLTILN 491
Qy 288 -TVEFCPQRQQTNAVTVYCGGHKGSAN-----VTVVNEPCVQVNISGA 333
Db 492 QTDICKHHANLCLNGRCIPTVSYRCECNMGYKQDANGDCIDVDECTSNPCTNGDCVNT 551
Qy 334 DWSYVCK-----PVEYSISVNSGDLVLDVVDQTLPSGVTVLEA 374
Db 552 PGSYCKCHAGFORTPTKQACIDIDEICQNGVLCKN--GRCVNSDGSFQICNAGFEL--T 608
Qy 375 PGGEIC-----CNKVVMRIKEMCPGETLQPKLVVK-----AQVPGRFT 412
Db 609 TDGKNVVDHDECITTNCLNGMCLNEDGSFKICKPGFVLAPNGRYCTDVDECTPGICM 668
Qy 413 NQVAVTSESN--CGTCTSCAETHTHWKGLAA-----THM---CVLDTNDPDCV---455
Db 669 NGHCINSEGSFRC---DCP-----PGLAVGMDGRVCDVTHMRSTCYGGIKKGVCVRPF 718
Qy 456 -GENTYRICVTN--RGSADTNVSLILKFSKELOPTASSG 493
Db 719 PGAVTKSECCANPDYGFGEPCQ-PCPAKNSAEFHGLCSSG 758

RESULT 13
DP87 DICI
ID DP87 DICI STANDARD; PRT; 555 AA.
AC Q04503:
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Prespore protein DP87 precursor.
GN DP87.
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxId=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94008533; PubMed=8404532;
RA Ozaki T., Nakao H., Orii H., Morio T., Takeuchi I., Tasaka M.;
RT "Developmental regulation of transcription of a novel
RL prespore-specific gene (dp87) in Dictyostellium discoideum.";
RL Development 117:1299-1308(1993).
CC -!- SUBCELLULAR LOCATION: STORED IN PRESPORE VACUOLES UNTIL IT IS
CC DISCHARGED INTO THE INTERSPACE OF SPORES DURING SPORE FORMATION.
CC -!- INDUCTION: BY EXOGENOUS CAMP, REPRESSED BY DIP.
CC -!- SIMILARITY: CONTAINS 7 PRESPORE MOTIFS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D13973; BAA03083.1; -.
DR HSSP: P01542; LCCN.
DR DictyDb: DD03013; -.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR003645; FOLN.
DR SMART: SM00001; EGF-like; 1.
DR SMART: SM00274; FOLN; 6.
KW Glycoprotein; Sporulation; Signal; Repeat.
```

```
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 555 PRESPORE PROTEIN DP87.
FT REPEAT 149 181 PRESPORE MOTIF 1.
FT REPEAT 176 188 PRESPORE MOTIF 2.
FT REPEAT 210 222 PRESPORE MOTIF 3.
FT REPEAT 241 253 PRESPORE MOTIF 4.
FT REPEAT 305 317 PRESPORE MOTIF 5.
FT REPEAT 333 345 PRESPORE MOTIF 6.
FT REPEAT 363 375 PRESPORE MOTIF 7.
FT DOMAIN 438 473 THR-RICH.
FT DOMAIN 476 555 ALA/SER-RICH.
FT CARBOHYD 103 103 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 555 AA; 58673 MW; 6D8A6DLC36F8DDD CRC64;

Query Match 3.8%; Score 113.5; DB 1; Length 555;
Best Local Similarity 17.4%; Pred. No. 0.36;
Matches 97; Conservative 66; Mismatches 215; Indels 179; Gaps 27;

Qy 75 CDKEFYPC-----EGRCPQVPAQOESCGRLYSVKVNDDCNVEICSVPEYATVGSYPPI 130
Db 54 CKKEFF-CVNRDRHNCDDLSCAKNTRTGEIFEIWS-----CKPHRDFVYHSP---102
Qy 131 ETLAIGKDCVDVVTIOQLPCEA-----EFVSSDPETTTPTSDGKLVMKIDRLGAGDKCK1 185
Db 103 -----NTTTCESLGCEARGMECEWVESPCYGTSCCPRIPCVGHGHHGKCDR 151
Qy 186 -----TVVKPLKEGC-----CFTATVCA-----205
Db 152 MRCEGFYCEEQGSACCVPHHDGCGNIQCPWGHYCVNEHGKCLVPHRPPRPVDPQCR 211
Qy 206 ---CPELRSYTKCGQPAICIKQEGPD---CACLRCPVCYKIEVVNTGSAIARNVTVDNPV 259
Db 212 NQCHPGYSCRVIKGCATCVRDARPHNLRCRGCEGSHCEVLEKHPVCVRNHRVPPHP 271
Qy 260 PDGYSHASQORVLSFNLGDMRPGDKKVFVEFCPQRQGITNVATVTCG-GHKCSANVT 318
Db 272 PP-----PQICGSVN-----CQPGYIC-----288
Qy 319 TVVN--EPCVQ-----VNISGADMSYVCKPVEYSISVNSPGDLVLDVVDQTLPSGVTV 371
Db 289 TIIINGHPTCIKRGDYLNCNTRCPHDYQCETISTINI VKSP-----KNDECKWHKCPPGSSC 344
Qy 372 LEAPGGEIC-CNKVVMRIKEM--CPGETLQFKLV-----VKAQVPGRFTNQVAVTSESN 422
Db 345 FNSRNGPHCLANNVFPQICKVTCQPTD-FSCKMIRGNPTCIKARPP-----VPPPH 394
Qy 423 CGTCTSCAETHTHWKGLAATHMCV-----LDTNDPICGENTVYRICV---TNRGSAD 473
Db 395 CSTCAELSSACNH-----VGMICIQVPSNCTNTRFPCPS---HPICIHPSITTAASTIA 445
Qy 474 TNVSLILKFSKELOPIASSGPKGTISGTVVVDALPKLGSKESEVFSVTLKGIAPGDAR 533
Db 446 TTASTV-----ATTTSATTAGTTGTTGGTSDSSAASSADSSAA--SSSPSSSA 495

Qy 534 GEAILSSDTLTSPVSDT 550
Db 496 ASSAASSEPPSSSAASS 512

RESULT 14
ID TENA_CHICK STANDARD; PRT; 1808 AA.
AC P10039; P13132; O73584; O73585;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tenascin precursor (TN) (Hexabrachion) (Cytotactin) (Neuronectin)
DE (GMEM) (J1) (Mitotendinous antigen) (Glioma-associated-extracellular
DE matrix antigen) (GP 150-225).
OS Gallus gallus (Chicken).
```

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Archosauaria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=90030407; PubMed=2478295;
 RA Spring J., Beck K., Chiquet-Ehrismann R.;
 RT "Two contrary functions of tenascin: dissection of the active sites
 by recombinant tenascin fragments.";
 RL Cell 59:325-334(1989).
 RN [2]
 RN SEQUENCE OF 27-722 FROM N.A., AND SEQUENCE OF 79-96.
 RC TISSUE=Fibroblast;
 RX MEDLINE=89030589; PubMed=2460335;
 RA Pearson C.A., Pearson D., Shibahara S., Hofsteenge J.,
 Chiquet-Ehrismann R.;
 RT "Tenascin: cDNA cloning and induction by TGF-beta.";
 RL EMBO J. 7:2977-2982(1988).
 RN [3]
 RN SEQUENCE OF 464-1018 AND 1412-1661 FROM N.A., AND SEQUENCE OF 852-868.
 RC TISSUE=Embryo;
 RX MEDLINE=88176910; PubMed=2451243;
 RA Jones F.S., Burgoon M.P., Hoffman S., Crossin K.L., Cunningham B.A.,
 Edelman G.M.;
 RT "A cDNA clone for cytotactin contains sequences similar to epidermal
 growth factor-like repeats and segments of fibronectin and
 fibrinogen";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:2186-2190(1988).
 CC -|- FUNCTION: SAM (SUBSTRATE-ADHESION MOLECULE) THAT APPEARS TO
 INHIBIT CELL MIGRATION. MAY PLAY A ROLE IN SUPPORTING THE GROWTH
 OF EPITHELIAL TUMORS. IS A LIGAND FOR INTEGRINS ALPHA-8/BETA-1,
 ALPHA-9/BETA-1, ALPHA-V/BETA-3 AND ALPHA-V/BETA-6.
 CC -|- SUBUNIT: HEXAMERIC. A HOMOTRIMER MAY BE FORMED IN THE TRIPLE
 COILED-COIL REGION AND MAY BE STABILIZED BY DISULFIDE RINGS AT
 BOTH ENDS. TWO OF SUCH HALF-HEXABRACHIONS MAY BE DISULFIDE LINKED
 WITHIN THE CENTRAL GLOBULE.
 CC -|- SUBCELLULAR LOCATION: Extracellular matrix.
 CC -|- ALTERNATIVE PRODUCTS: 3 ISOFORMS: 230 KDA (SHOWN HERE), 200 KDA
 AND 190 KDA; ARE PRODUCED BY ALTERNATIVE SPLICING. THEY ARE
 PRODUCED IN A TISSUE- AND TIME-SPECIFIC MANNER DURING DEVELOPMENT.
 CC -|- INDUCTION: BY TGF-BETA.
 CC -|- SIMILARITY: CONTAINS 13.5 EGF-LIKE DOMAINS.
 CC -|- SIMILARITY: CONTAINS 11 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -|- SIMILARITY: CONTAINS A DOMAIN RELATED TO THE C-TERMINAL PART OF
 THE BETA AND GAMMA CHAINS OF FIBRINOGEN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M23121; AAA49086.1; -;
 DR EMBL; X08031; CAB40811.1; -;
 DR EMBL; X08030; CAA30824.1; ALT-TERM.
 DR EMBL; J03641; AAA48748.1; ALT-SEQ.
 DR EMBL; M20816; AAA48749.1; ALT-SEQ.
 DR PIR; A30903; A30903.
 DR PIR; A31930; A31930.
 DR PIR; A33379; A33379.
 DR PIR; B33379; B33379.
 DR PIR; C33379; C33379.
 DR PIR; S01292; S01292.
 DR HSSP; P24821; 1TEN.
 DR InterPro; IPR000361; EGF-like.
 DR InterPro; IPR003961; FN.III.
 DR InterPro; IPR002181; Fibrinogen_C.
 DR Pfam; PF00008; EGF; 13.
 DR Pfam; PF00147; fibrinogen_C; 1.

DR Pfam; PF00041; fn3; 11.
 DR SMART; SM00181; EGF; 10.
 DR SMART; SM00001; EGF_Like; 1.
 DR SMART; SM00186; FBG; 1.
 DR SMART; SM00060; FN3; 10.
 DR PROSITE; PS00022; EGF_1; 14.
 DR PROSITE; PS01186; EGF_2; 14.
 KW Glycoprotein; Cell adhesion;
 KW Extracellular matrix; Signal; Alternative splicing.
 FT SIGNAL 1 22
 FT PROPEP 23 33
 FT CHAIN 34 1808
 FT DOMAIN 118 142
 FT DOMAIN 176 188
 FT DOMAIN 188 219
 FT DOMAIN 219 250
 FT DOMAIN 250 281
 FT DOMAIN 281 312
 FT DOMAIN 312 343
 FT DOMAIN 343 374
 FT DOMAIN 374 405
 FT DOMAIN 405 436
 FT DOMAIN 436 467
 FT DOMAIN 467 498
 FT DOMAIN 498 529
 FT DOMAIN 529 560
 FT DOMAIN 560 591
 FT DOMAIN 591 680
 FT DOMAIN 681 771
 FT DOMAIN 772 862
 FT DOMAIN 863 954
 FT DOMAIN 955 1042
 FT DOMAIN 1043 1133
 FT DOMAIN 1134 1224
 FT DOMAIN 1225 1315
 FT DOMAIN 1316 1404
 FT DOMAIN 1405 1492
 FT DOMAIN 1493 1580
 FT DOMAIN 1589 1808
 FT DISULFID 64 64
 FT DISULFID 192 202
 FT DISULFID 196 207
 FT DISULFID 209 218
 FT DISULFID 223 233
 FT DISULFID 227 238
 FT DISULFID 240 249
 FT DISULFID 254 264
 FT DISULFID 258 269
 FT DISULFID 271 280
 FT DISULFID 285 295
 FT DISULFID 289 300
 FT DISULFID 302 311
 FT DISULFID 316 326
 FT DISULFID 320 331
 FT DISULFID 333 342
 FT DISULFID 347 357
 FT DISULFID 351 362
 FT DISULFID 364 373
 FT DISULFID 378 388
 FT DISULFID 382 393
 FT DISULFID 395 404
 FT DISULFID 409 419
 FT DISULFID 413 424
 FT DISULFID 426 435
 FT DISULFID 440 450
 FT DISULFID 444 455
 FT DISULFID 457 466
 FT DISULFID 471 481
 FT DISULFID 475 486
 FT DISULFID 488 497
 FT DISULFID 502 512
 FT DISULFID 506 517
 FT DISULFID 519 528

TENASCIN.

COILED COIL (POTENTIAL).

EGF-LIKE 1 (INCOMPLETE).

EGF-LIKE 2.

EGF-LIKE 3.

EGF-LIKE 4.

EGF-LIKE 5.

EGF-LIKE 6.

EGF-LIKE 7.

EGF-LIKE 8.

EGF-LIKE 9.

EGF-LIKE 10.

EGF-LIKE 11.

EGF-LIKE 12.

EGF-LIKE 13.

EGF-LIKE 14.

FIBRONECTIN TYPE-III 1.

FIBRONECTIN TYPE-III 2.

FIBRONECTIN TYPE-III 3.

FIBRONECTIN TYPE-III 4.

FIBRONECTIN TYPE-III 5.

FIBRONECTIN TYPE-III 6.

FIBRONECTIN TYPE-III 7.

FIBRONECTIN TYPE-III 8.

FIBRONECTIN TYPE-III 9.

FIBRONECTIN TYPE-III 10.

FIBRONECTIN TYPE-III 11.

FIBRINOGEN BETA/GAMMA.

INTERCHAIN (POTENTIAL).

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

```
FT DISULFID 533 BY SIMILARITY.
FT DISULFID 537 BY SIMILARITY.
FT DISULFID 550 BY SIMILARITY.
FT DISULFID 564 BY SIMILARITY.
FT DISULFID 568 BY SIMILARITY.
FT DISULFID 581 BY SIMILARITY.
FT CARBOHYD 38 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 168 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 328 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 603 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 643 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 751 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 759 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1050 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1090 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1101 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1112 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1153 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1183 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1416 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1736 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1769 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 1043 MISSING (IN 200 KDA ISOFORM).
FT VARSPLIC 1043 MISSING (IN 190 KDA ISOFORM).
FT CONFLICT 182 W -> R (IN REF. 2).
FT CONFLICT 563 SCPNDCNNV -> PAPMTATTW (IN REF. 3).
FT CONFLICT 598 E -> G (IN REF. 3).
FT CONFLICT 838 T -> TEX (IN REF. 3).
FT CONFLICT 886 N -> F (IN REF. 3).
FT SEQUENCE 1808 AA; 198858 MW; B924A06CF9EFD6DE CRC64;

Query Match
Best Local Similarity 3.8%; Score 113.5; DB 1; Length 1808;
Matches 127; Conservative 64; Mismatches 200; Indels 213; Gaps 37;

Qy 82 CEEGRCPVEAQECYGLYSKVNDCNVEIC-----VITQQLPCEAEFVSDDPETTPSDGKLWKIDRL 177
Db 357 CENGCLC-----VCHEGF-----VGDDCSOKRCPKDCNNRGHCVDGRCVCHGYLG--- 401
Qy 129 PTEILAIGKKDCVDV-----VITQQLPCEAEFVSDDPETTPSDGKLWKIDRL 177
Db 402 -----EDCGELRCPNDCNHRGRCINGQCVDGDFIGED-----CGEL 438
Qy 178 GAGDKCKITVWVKPEKGCCTAATVC-----ACPELRSTYKCGOAPICIKOE----- 225
Db 439 RCPNDCH-----NRCRCVNGQCECHGEFGICEDCGELRCPNDCNCSHGRCVNGQCVCDE 490
Qy 236 ---GPDCACLRCP-VCYKIEVVNTGSAIARNVTVDNVPDGY-----SHASG 268
Db 491 GYTGEDCGLRCPNDCN-----NRCRCVNGRCVCDN-----GFMGEDCGLSCPNDCHQHG 541
Qy 269 QRVLFSNLGDMRPGDKVFTVEFCPPQRGQITNVAITVTCGGHKCSANVTTVVNEPCV-Q 327
Db 542 RCV-----DGRVCVCHGEFTGEDCRERS-----C-PNDCN-NVGRCVGRCVCE 582
Qy 328 VNISADNSYVCKPVEYSISVSNPDGLV-----HDVVI-----QDTL 365
Db 583 EGYMGIDCSDVSPTE--LTVTNVTDKTVNLEWKHENLVNLYVTVPVTS SSGGLDLQFTV 640
Qy 366 P---SGVTVLE-APGGE-----ICCNK-----VWRIKEMCPG-ETLQFKLVVKAQVPG 409
Db 641 PGNQTSATIHLEPQVEFIRFVAILKNKKSIPVSARVATYLPAPGLKFKFSVRETSVQV 700
Qy 410 R-----FTNQVAVTSBSNGCTCTSCAETTHWKGLA-----ATHMCVLID 448
Db 701 EWDPLSISDGMELVFRNMQKDDNGDITSSLKREPTSYMQGLAPQQYNNVSLHIVKNN 760
Qy 449 TNDPICVGENTVY--RICVTNRGSED--TNVSLIKFSKELQPIASSGPTKG--TISGNT 503
Db 761 TRGP---GLSRVITTKLDAPSQIEAKVDVTDITATLTWTSKPLAEIEGIELTYGPKDVPDGR 817
```

```
Qy 504 VVFDALP-----KLGs-KESVFSVTL---KGIAPGDARCEAILSDTLTSP-----VSD 549
Db 818 TTIDLSEDNQYISGNLPRHTTEVTLLSRGDMESDPAKEVEVTD--LDAPRNLKRVSQ 875
Qy 550 TENT 553
Db 876 TDNS 879

RESULT 15
FBN2_MOUSE
ID FBN2_MOUSE STANDARD: PRT; 2907 AA.
AC O61555; O63957;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Fibrillin 2 precursor.
GN FBN2 OR FBN-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95263670; PubMed=7744963;
RA Zhang H., Hu W., Ramirez F.;
RT "Developmental expression of fibrillin genes suggests heterogeneity
of extracellular microfibrils.";
RL J. Cell Biol. 129:1165-1176(1995).
RN [2]
RP SEQUENCE OF 210-317 FROM N.A.
RX MEDLINE=94140368; PubMed=8307578;
RA Li X., Pereira L., Zhang H., Sanguinetti C., Ramirez F., Bonadio J.,
RA Francke U.;
RT "Fibrillin genes map to regions of conserved mouse/human synteny on
mouse chromosomes 2 and 18.";
RL Genomics 18:667-672(1993).
CC -1- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE
LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
CC -1- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING
EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; L39790; AAA74908.1; -
CC EMBL; S65359; AAC60685.1; -
CC HSP; P35555; 1EMN.
CC MGD; MGI:95490; Fbn2.
CC InterPro; IPR000152; Asx_hydroxyl.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR001438; EGF_II.
CC InterPro; IPR002212; Tn.
CC InterPro; IPR000822; Znf-C2H2.
CC Pfam; PF00008; EGF; 46.
CC Pfam; PF00683; TB; 9.
CC PRINTS; PR00010; EGFBL00D.
CC SMART; SM00179; EGF_CA; 43.
CC SMART; SM00001; EGF_Like; 3.
CC PROSITE; PS00010; ASX_HYDROXYL; 43.
CC PROSITE; PS00022; EGF_1; 2.
CC PROSITE; PS01186; EGF_2; 36.
CC PROSITE; PS01187; EGF_CA; 43.
CC Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
Repeat; Signal; Multigene family.
KW
```

FT	SIGNAL	1	28	POTENTIAL	FT	DISULFID	514	526	BY SIMILARITY.
FT	CHAIN	29	2907	FIBRILLIN 2.	FT	DISULFID	532	542	BY SIMILARITY.
FT	EGF-LIKE 1, NON-CALCIUM BINDING.	111	142	EGF-LIKE 1, NON-CALCIUM BINDING.	FT	DISULFID	537	551	BY SIMILARITY.
FT	EGF-LIKE 2, NON-CALCIUM BINDING.	145	176	EGF-LIKE 2, NON-CALCIUM BINDING.	FT	DISULFID	553	566	BY SIMILARITY.
FT	EGF-LIKE 3, NON-CALCIUM BINDING.	176	208	EGF-LIKE 3, NON-CALCIUM BINDING.	FT	DISULFID	572	584	BY SIMILARITY.
FT	EGF-LIKE 4, CALCIUM-BINDING.	216	317	EGF-LIKE 4, CALCIUM-BINDING.	FT	DISULFID	579	593	BY SIMILARITY.
FT	EGF-LIKE 5, CALCIUM-BINDING.	318	359	EGF-LIKE 5, CALCIUM-BINDING.	FT	DISULFID	595	608	BY SIMILARITY.
FT	TGFBP 1.	360	426	TGFBP 1.	FT	DISULFID	614	625	BY SIMILARITY.
FT	EGF-LIKE 6, NON-CALCIUM BINDING.	487	527	EGF-LIKE 6, NON-CALCIUM BINDING.	FT	DISULFID	620	634	BY SIMILARITY.
FT	EGF-LIKE 7, CALCIUM-BINDING.	528	567	EGF-LIKE 7, CALCIUM-BINDING.	FT	DISULFID	636	649	BY SIMILARITY.
FT	EGF-LIKE 8, CALCIUM-BINDING.	568	609	EGF-LIKE 8, CALCIUM-BINDING.	FT	DISULFID	655	666	BY SIMILARITY.
FT	EGF-LIKE 9, CALCIUM-BINDING.	610	650	EGF-LIKE 9, CALCIUM-BINDING.	FT	DISULFID	661	675	BY SIMILARITY.
FT	EGF-LIKE 10, CALCIUM-BINDING.	651	691	EGF-LIKE 10, CALCIUM-BINDING.	FT	DISULFID	677	690	BY SIMILARITY.
FT	TGFBP 2.	692	760	TGFBP 2.	FT	DISULFID	765	777	BY SIMILARITY.
FT	EGF-LIKE 11, CALCIUM-BINDING.	761	802	EGF-LIKE 11, CALCIUM-BINDING.	FT	DISULFID	772	786	BY SIMILARITY.
FT	EGF-LIKE 12, CALCIUM-BINDING.	803	844	EGF-LIKE 12, CALCIUM-BINDING.	FT	DISULFID	788	801	BY SIMILARITY.
FT	EGF-LIKE 13, CALCIUM-BINDING.	845	883	EGF-LIKE 13, CALCIUM-BINDING.	FT	DISULFID	807	819	BY SIMILARITY.
FT	EGF-LIKE 14, CALCIUM-BINDING.	889	989	EGF-LIKE 14, CALCIUM-BINDING.	FT	DISULFID	814	828	BY SIMILARITY.
FT	TGFBP 3.	1065	1065	TGFBP 3.	FT	DISULFID	830	843	BY SIMILARITY.
FT	EGF-LIKE 15, CALCIUM-BINDING.	1107	1107	EGF-LIKE 15, CALCIUM-BINDING.	FT	DISULFID	849	859	BY SIMILARITY.
FT	EGF-LIKE 16, CALCIUM-BINDING.	1108	1150	EGF-LIKE 16, CALCIUM-BINDING.	FT	DISULFID	854	868	BY SIMILARITY.
FT	EGF-LIKE 17, CALCIUM-BINDING.	1151	1192	EGF-LIKE 17, CALCIUM-BINDING.	FT	DISULFID	870	883	BY SIMILARITY.
FT	EGF-LIKE 18, CALCIUM-BINDING.	1193	1234	EGF-LIKE 18, CALCIUM-BINDING.	FT	DISULFID	952	964	BY SIMILARITY.
FT	EGF-LIKE 19, CALCIUM-BINDING.	1235	1275	EGF-LIKE 19, CALCIUM-BINDING.	FT	DISULFID	959	973	BY SIMILARITY.
FT	EGF-LIKE 20, CALCIUM-BINDING.	1317	1317	EGF-LIKE 20, CALCIUM-BINDING.	FT	DISULFID	975	988	BY SIMILARITY.
FT	EGF-LIKE 21, CALCIUM-BINDING.	1318	1359	EGF-LIKE 21, CALCIUM-BINDING.	FT	DISULFID	1070	1082	BY SIMILARITY.
FT	EGF-LIKE 22, CALCIUM-BINDING.	1360	1400	EGF-LIKE 22, CALCIUM-BINDING.	FT	DISULFID	1077	1091	BY SIMILARITY.
FT	EGF-LIKE 23, CALCIUM-BINDING.	1401	1441	EGF-LIKE 23, CALCIUM-BINDING.	FT	DISULFID	1093	1106	BY SIMILARITY.
FT	EGF-LIKE 24, CALCIUM-BINDING.	1442	1483	EGF-LIKE 24, CALCIUM-BINDING.	FT	DISULFID	1112	1124	BY SIMILARITY.
FT	EGF-LIKE 25, CALCIUM-BINDING.	1484	1524	EGF-LIKE 25, CALCIUM-BINDING.	FT	DISULFID	1119	1133	BY SIMILARITY.
FT	EGF-LIKE 26, CALCIUM-BINDING.	1525	1565	EGF-LIKE 26, CALCIUM-BINDING.	FT	DISULFID	1135	1149	BY SIMILARITY.
FT	TGFBP 4.	1566	1642	TGFBP 4.	FT	DISULFID	1155	1167	BY SIMILARITY.
FT	EGF-LIKE 27, CALCIUM-BINDING.	1643	1684	EGF-LIKE 27, CALCIUM-BINDING.	FT	DISULFID	1162	1176	BY SIMILARITY.
FT	EGF-LIKE 28, CALCIUM-BINDING.	1685	1726	EGF-LIKE 28, CALCIUM-BINDING.	FT	DISULFID	1178	1191	BY SIMILARITY.
FT	TGFBP 5.	1727	1800	TGFBP 5.	FT	DISULFID	1197	1209	BY SIMILARITY.
FT	EGF-LIKE 29, CALCIUM-BINDING.	1801	1842	EGF-LIKE 29, CALCIUM-BINDING.	FT	DISULFID	1204	1218	BY SIMILARITY.
FT	EGF-LIKE 30, CALCIUM-BINDING.	1843	1884	EGF-LIKE 30, CALCIUM-BINDING.	FT	DISULFID	1220	1233	BY SIMILARITY.
FT	EGF-LIKE 31, CALCIUM-BINDING.	1885	1926	EGF-LIKE 31, CALCIUM-BINDING.	FT	DISULFID	1239	1250	BY SIMILARITY.
FT	EGF-LIKE 32, CALCIUM-BINDING.	1927	1965	EGF-LIKE 32, CALCIUM-BINDING.	FT	DISULFID	1246	1259	BY SIMILARITY.
FT	EGF-LIKE 33, CALCIUM-BINDING.	1966	2008	EGF-LIKE 33, CALCIUM-BINDING.	FT	DISULFID	1261	1274	BY SIMILARITY.
FT	EGF-LIKE 34, CALCIUM-BINDING.	2009	2048	EGF-LIKE 34, CALCIUM-BINDING.	FT	DISULFID	1280	1292	BY SIMILARITY.
FT	EGF-LIKE 35, CALCIUM-BINDING.	2049	2090	EGF-LIKE 35, CALCIUM-BINDING.	FT	DISULFID	1287	1301	BY SIMILARITY.
FT	TGFBP 6.	2091	2163	TGFBP 6.	FT	DISULFID	1303	1316	BY SIMILARITY.
FT	EGF-LIKE 36, CALCIUM-BINDING.	2164	2205	EGF-LIKE 36, CALCIUM-BINDING.	FT	DISULFID	1322	1334	BY SIMILARITY.
FT	EGF-LIKE 37, CALCIUM-BINDING.	2206	2245	EGF-LIKE 37, CALCIUM-BINDING.	FT	DISULFID	1329	1343	BY SIMILARITY.
FT	EGF-LIKE 38, CALCIUM-BINDING.	2246	2286	EGF-LIKE 38, CALCIUM-BINDING.	FT	DISULFID	1345	1358	BY SIMILARITY.
FT	EGF-LIKE 39, CALCIUM-BINDING.	2287	2330	EGF-LIKE 39, CALCIUM-BINDING.	FT	DISULFID	1364	1377	BY SIMILARITY.
FT	EGF-LIKE 40, CALCIUM-BINDING.	2331	2372	EGF-LIKE 40, CALCIUM-BINDING.	FT	DISULFID	1371	1386	BY SIMILARITY.
FT	TGFBP 7.	2373	2441	TGFBP 7.	FT	DISULFID	1388	1399	BY SIMILARITY.
FT	EGF-LIKE 41, CALCIUM-BINDING.	2442	2483	EGF-LIKE 41, CALCIUM-BINDING.	FT	DISULFID	1405	1418	BY SIMILARITY.
FT	EGF-LIKE 42, CALCIUM-BINDING.	2484	2524	EGF-LIKE 42, CALCIUM-BINDING.	FT	DISULFID	1412	1427	BY SIMILARITY.
FT	EGF-LIKE 43, CALCIUM-BINDING.	2525	2563	EGF-LIKE 43, CALCIUM-BINDING.	FT	DISULFID	1429	1440	BY SIMILARITY.
FT	EGF-LIKE 44, CALCIUM-BINDING.	2564	2606	EGF-LIKE 44, CALCIUM-BINDING.	FT	DISULFID	1446	1458	BY SIMILARITY.
FT	EGF-LIKE 45, CALCIUM-BINDING.	2607	2646	EGF-LIKE 45, CALCIUM-BINDING.	FT	DISULFID	1453	1467	BY SIMILARITY.
FT	EGF-LIKE 46, CALCIUM-BINDING.	2647	2687	EGF-LIKE 46, CALCIUM-BINDING.	FT	DISULFID	1469	1482	BY SIMILARITY.
FT	EGF-LIKE 47, CALCIUM-BINDING.	2688	2727	EGF-LIKE 47, CALCIUM-BINDING.	FT	DISULFID	1488	1499	BY SIMILARITY.
FT	BY SIMILARITY.	115	124	BY SIMILARITY.	FT	DISULFID	1494	1508	BY SIMILARITY.
FT	BY SIMILARITY.	119	130	BY SIMILARITY.	FT	DISULFID	1510	1523	BY SIMILARITY.
FT	BY SIMILARITY.	132	141	BY SIMILARITY.	FT	DISULFID	1529	1540	BY SIMILARITY.
FT	BY SIMILARITY.	149	159	BY SIMILARITY.	FT	DISULFID	1535	1549	BY SIMILARITY.
FT	BY SIMILARITY.	153	164	BY SIMILARITY.	FT	DISULFID	1551	1564	BY SIMILARITY.
FT	BY SIMILARITY.	166	175	BY SIMILARITY.	FT	DISULFID	1647	1659	BY SIMILARITY.
FT	BY SIMILARITY.	180	190	BY SIMILARITY.	FT	DISULFID	1654	1668	BY SIMILARITY.
FT	BY SIMILARITY.	184	196	BY SIMILARITY.	FT	DISULFID	1670	1683	BY SIMILARITY.
FT	BY SIMILARITY.	198	207	BY SIMILARITY.	FT	DISULFID	1689	1701	BY SIMILARITY.
FT	BY SIMILARITY.	280	292	BY SIMILARITY.	FT	DISULFID			
FT	BY SIMILARITY.	287	301	BY SIMILARITY.	FT	DISULFID			
FT	BY SIMILARITY.	303	316	BY SIMILARITY.	FT	DISULFID			
FT	BY SIMILARITY.	322	334	BY SIMILARITY.	FT	DISULFID			
FT	BY SIMILARITY.	329	343	BY SIMILARITY.	FT	DISULFID			
FT	BY SIMILARITY.	345	358	BY SIMILARITY.	FT	DISULFID			
FT	BY SIMILARITY.	491	503	BY SIMILARITY.	FT	DISULFID			
FT	BY SIMILARITY.	498	512	BY SIMILARITY.	FT	DISULFID			

Query Match 3.8%; Score 113; DB 1; Length 2907;
Best Local Similarity 18.5%; Pred. No. 2.6;
Matches 107; Conservative
62 NKQPVQKSRGAFCDKEFYPCCEGR----COPVEAQOQESCY-GRLYSVKVNDDCNVEIC 115
221 NNQMCQGLTGIVCTKTLCCCATIGRAWHPCEMCPAQPCRPFPINRTGACQDVDEC 280

Search completed: May 25, 2002, 22:24:52
Job time: 490 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

DM protein - protein search, using sw model

Run on: May 25, 2002, 22:15:37 ; Search time 95.87 Seconds
(without alignments)
1003.288 Million cell updates/sec

Title: US-09-523-647-2
Perfect score: 2951
Sequence: 1 MSKLIRRVTVLTALTSMA...ILSDTLTSPVSDTENTHYV 556

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL19.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2523.5	85.5	557	Q9AIS7	Q9ais7 chlamydophi
2	2488	84.3	558	P94664	P94664 chlamydophi
3	2418.5	82.0	534	Q9ZAW1	Q9zaw1 chlamydophi
4	2323.5	78.7	534	Q9ZAW0	Q9zaw0 chlamydophi
5	2176	73.7	554	Q9PJY0	Q9pjy0 chlamydia m
6	2157.5	73.1	547	Q933I7	Q933i7 chlamydia t
7	2127	72.1	395	Q956B3	Q956b3 chlamydia p
8	2008.5	68.1	511	Q93FV8	Q93fv8 chlamydia t
9	1917	65.0	395	Q9X4I2	Q9x4i2 chlamydophi
10	1833	62.1	395	Q9X4I1	Q9x4i1 chlamydophi
11	956	32.4	183	Q46164	Q46164 chlamydia p
12	920	31.2	176	Q9AG10	Q9ag10 chlamydia p
13	685	23.2	182	Q9AF83	Q9af83 chlamydophi
14	674	22.8	128	Q9RB53	Q9rb53 chlamydia p
15	669	22.7	178	Q46163	Q46163 chlamydophi
16	659	22.3	178	Q46165	Q46165 chlamydophi

17	578	19.6	157	2	Q9AF85	Q9af85 chlamydophi
18	578	19.6	157	2	Q9AF84	Q9af84 chlamydophi
19	509	17.2	173	2	Q9AF82	Q9af82 chlamydia s
20	151.5	5.1	1474	17	O27146	O27146 methanother
21	135	4.6	2233	5	Q94711	Q94711 paramecium
22	130.5	4.4	756	16	Q9RZS5	Q9rzs5 deinococcus
23	129.5	4.4	667	5	Q95WU1	Q95wu1 giardia lam
24	129.5	4.4	757	5	Q9VZF2	Q9vzf2 drosophila
25	129	4.4	556	5	Q9NGZ3	Q9ngz3 giardia lam
26	128.5	4.4	1404	5	Q9VB65	Q9vb65 drosophila
27	127.5	4.3	998	16	Q9RWI2	Q9rwi2 deinococcus
28	127.5	4.3	2327	13	Q9IBG7	Q9ibg7 xenopus lae
29	127	4.3	693	5	O01600	O01600 caenorhabdi
30	125.5	4.3	3198	5	Q9U8G8	Q9u8g8 manduca sex
31	124.5	4.2	739	5	Q9GS24	Q9gs24 giardia lam
32	124	4.2	2440	5	Q9GP97	Q9gp97 caenorhabdi
33	123.5	4.2	765	5	Q9NL50	Q9nl50 sarcophaga
34	123.5	4.2	2174	5	Q9GQR0	Q9gqr0 drosophila
35	123.5	4.2	2906	11	Q9WUH9	Q9wuh9 rattus norv
36	122.5	4.2	950	13	Q90Z44	Q90z44 gallus gall
37	122	4.1	491	16	Q9RY09	Q9ry09 deinococcus
38	122	4.1	955	4	Q96DN2	Q96dn2 homo sapien
39	122	4.1	2910	11	O55225	O55225 mus musculu
40	122	4.1	2931	5	Q9W2C6	Q9w2c6 drosophila
41	121.5	4.1	2809	4	Q96JP8	Q96jp8 homo sapien
42	121	4.1	641	17	O58295	O58295 pyrococcus
43	121	4.1	1546	4	Q9NS27	Q9ns27 homo sapien
44	121	4.1	1551	4	O75445	O75445 homo sapien
45	120.5	4.1	354	5	Q9XV21	Q9xv21 caenorhabdi

ALIGNMENTS

RESULT 1

Q9AIS7	Q9AIS7	PRELIMINARY;	PRT;	557 AA.
ID	Q9AIS7			
AC	Q9AIS7;			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-OCT-2001 (TREMBLrel. 18, Last annotation update)			
DE	60 KDA CYSTEINE-RICH MEMBRANE COMPLEX PROTEIN.			
GN	CMCB.			
OS	Chlamydophila abortus.			
OC	Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.			
OX	NCBI_TaxID=893555;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=EBA;			
RX	MEDLINE=21078680; PubMed=11211261;			
RA	Bush R.M., Everett K.D.;			
RT	"Molecular evolution of the Chlamydiaceae.";			
RL	Int. J. Syst. Evol. Microbiol. 51:203-220(2001).			
DR	EMBL; AF240773; AAG60550.1; -			
DR	InterPro: IPR003506; Chlam.OMP6.			
DR	PRINTS: PR01336; CHLAMIDIOM6.			
SQ	SEQUENCE 557 AA; 59761 MW; 817A20B071AD9609 CRC64;			

Query Match	85.5%;	Score 2523.5;	DB 2;	Length 557;
Best Local Similarity	84.4%;	Pred. No. 3.6e-198;		
Matches 470;	Conservative 33;	Mismatches 53;	Indels 1;	Gaps 1;
1	MSKLIRRVTVLTALTSMA	SCFASGGTAAVAESLI	TKIVASAETKPAVP-PMTAKKVR	LV 59
1	MSKLIRRVTVLTALTSMA	SSFGSKTEAAAAA	SLATRFIASTENSNDNV	LQATAKKVRFG 60
60	RNRNQPEQKSRGAFCDKE	FPYPCBEGRCQV	EAQQAQSCYGR	LYSVKVNDDCNVEICQSV 119
61	RNRNQPEQKHQAFCDKE	FPYPCBEGRCQ	SCYCKMYCVRVNDDC	NCNVEISQAVP 120
120	EYATVGSPPYIEIATIGK	KDCVDVVTIQQLPC	BAEFVSSDPETPT	SDGKLVKIDRLGA 179

```

Db 121 EYATVGSPPYIEILAVGKKDCVNVVITQOLPCEVEFVSSDPATPTSDSKLIWTIDCLGQ 180
QY 180 GDKCKITVWVKPLKEGCCCTAATVACAPELRSYTKCGQPAICIKOEGPDCACLRCPVCYK 239
Db 181 QGERCKITVWVKPLKEGCCCTAATVACAPELRSYTKCGQPAICIKOEGPDCACLRCPVCYK 240
Db 181 GEKCKITVWVKPLKEGCCCTAATVACAPELRSYTKCGQPAICIKOEGPDCACLRCPVCYK 240
QY 240 IEVNTGSAIARNVTVDNVPDGYSHASQORVLSFNLDGMRPDKKVFTEFCPQRRGOI 299
Db 241 IEVNTGSAIARNVTVDNVPDGYSHASQORVLSFNLDGMRPDKKVFTEFCPQRRGOI 299
QY 300 TNVATVYCGGHKCSANVTTVVNEPCVQVNIISGADMSYVCKPVEYSISVSNPGDLVLHDV 359
Db 301 TNVATVYCGGHKCSANVTTVVNEPCVQVNIISGADMSYVCKPVEYSISVSNPGDLVLHDV 359
QY 360 VIQDTLPSPGVTLAPGGEICCNKVVWRIKEMCPGETLQFKLVKVAQVPGRETNOVAVTS 419
Db 361 VVEDTVPAGATILEAEGAEICCNKAVWCIKEMCPGETLQFKLVKVAQVPGRETNOVAVTS 419
QY 420 ESNCGTCTSCAETTHWKGLAATHMCVLDTDNDPICVGENTVYRICVTVNRGSAEDTNVSLI 479
Db 421 NSDCGTCTSCAETTHWKGLAATHMCVLDTDNDPICVGENTVYRICVTVNRGSAEDTNVSLI 480
QY 480 LKFSKELQPIASSGPTKGTISGNTVVFDPALPKLGSKEVESVTLKGIAPGDARGEAIL 539
Db 481 LKFSKELQPIASSGPTKGTISGNTVVFDPALPKLGSKEVESVTLKGIAPGDARGEAIL 539
QY 540 SDTLTSPVSDTENTHVV 556
Db 541 SDTLTVPVADTENTHVV 557

```

```

RESULT 2
P94664
ID Q92AW1 PRELIMINARY; PRT; 558 AA.
AC P94664;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE OUTER MEMBRANE PROTEIN 2.
GN OMP2
OS Chlamydomophila caviae.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OX NCBI_TaxID=83557;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97075924; PubMed=8918247;
RX MEDLINE=97075924; PubMed=8918247;
RA Hsia R.C., Bavoil P.M.;
RT "Sequence analysis of the omp2 region of Chlamydia psittaci strain GPIC: structural and functional implications.";
RL Gene 176:155-162(1996).
DR EMBL; U41759; AAB41143.1; -
DR InterPro; IPR003506; Chlam_Omp6.
DR PRINTS; PR01336; CHLAMIDIAOM6.
SQ SEQUENCE 558 AA; 60213 MW; 452E2BD79078935C CRC64;

```

```

Query Match 84.3%; Score 2488; DB 2; Length 558;
Best Local Similarity 82.1%; Pred. No. 2.9e-195;
Matches 458; Conservative 44; Mismatches 54; Indels 2; Gaps 2;
QY 1 MSLKTRRVTVLALTSMAFCFASGGIEAAVAESLITKIVASAEETKPAPVM-TAKKRVLV 59
Db 1 MSLKTRRVTVLALTSMAFCFASGGIEAAVAESLITKIVASAEETKPAPVM-TAKKRVLV 59
QY 60 RNKOPVQKSGARCDKEFYCEGRQ-PVEAQOESCYGRLYSVKVNDDCNVEICQSV 118
Db 61 RNKOPVQKSGARCDKEFYCEGRQ-PVEAQOESCYGRLYSVKVNDDCNVEICQSV 118
QY 119 PEYATVGSPPYIEILAVGKKDCVNVVITQOLPCEVEFVSSDPATPTSDSKLIWTIDCLG 178
Db 121 PEYATVGSPPYIEILAVGKKDCVNVVITQOLPCEVEFVSSDPATPTSDSKLIWTIDCLG 180

```

```

QY 179 AGDKCKITVWVKPLKEGCCCTAATVACAPELRSYTKCGQPAICIKOEGPDCACLRCPVCY 238
Db 181 QGERCKITVWVKPLKEGCCCTAATVACAPELRSYTKCGQPAICIKOEGPDCACLRCPVCY 240
QY 239 KIEVNTGSAIARNVTVDNVPDGYSHASQORVLSFNLDGMRPDKKVFTEFCPQRRGO 298
Db 241 KIEVNTGSAIARNVTVDNVPDGYSHASQORVLSFNLDGMRPDKKVFTEFCPQRRGO 298
QY 299 ITNVATVYCGGHKCSANVTTVVNEPCVQVNIISGADMSYVCKPVEYSISVSNPGDLVLHD 358
Db 301 ITNVATVYCGGHKCSANVTTVVNEPCVQVNIISGADMSYVCKPVEYSISVSNPGDLVLHD 358
QY 359 VIQDTLPSPGVTLAPGGEICCNKVVWRIKEMCPGETLQFKLVKVAQVPGRETNOVAVT 418
Db 361 VVEDTVPAGATILEAEGAEICCNKAVWCIKEMCPGETLQFKLVKVAQVPGRETNOVAVT 418
QY 419 SESNCGTCTSCAETTHWKGLAATHMCVLDTDNDPICVGENTVYRICVTVNRGSAEDTNVSL 478
Db 421 TNSDCGTCTSCAETTHWKGLAATHMCVLDTDNDPICVGENTVYRICVTVNRGSAEDTNVSL 480
QY 479 ILKFSKELQPIASSGPTKGTISGNTVVFDPALPKLGSKEVESVTLKGIAPGDARGEAIL 538
Db 481 ILKFSKELQPIASSGPTKGTISGNTVVFDPALPKLGSKEVESVTLKGIAPGDARGEAIL 538
QY 539 SDTLTSPVSDTENTHVV 556
Db 541 SDTLTVPVADTENTHVV 558

```

```

RESULT 3
Q92AW1
ID Q92AW1 PRELIMINARY; PRT; 534 AA.
AC Q92AW1;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CYSTEIN-RICH OUTER MEMBRANE PROTEIN OMP-2 (FRAGMENT).
OS Chlamydomophila abortus.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OX NCBI_TaxID=83555;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S26/3;
RA Sheehy N., Markey B., Quinn P.J.;
RT "Sequence analysis of C. psittaci and C. pecorum 60 kDa genes.";
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U76760; AAD09597.1; -
DR InterPro; IPR003506; Chlam_Omp6.
DR PRINTS; PR01336; CHLAMIDIAOM6.
FT NON_TER 1
FT NON_TER 534
SQ SEQUENCE 534 AA; 57137 MW; 2789820B041E92BE CRC64;

```

```

Query Match 82.0%; Score 2418.5; DB 2; Length 534;
Best Local Similarity 84.1%; Pred. NO. 1.3e-189;
Matches 449; Conservative 32; Mismatches 52; Indels 1; Gaps 1;
QY 10 TVLALTSMAFCFASGGIEAAVAESLITKIVASAEETKPAPV-PMTAKKRVLRNKPQVEQ 68
Db 1 TVLALTSMAFCFASGGIEAAVAESLITKIVASAEETKPAPV-PMTAKKRVLRNKPQVEQ 68
QY 69 KSRGAFCDKEFYCEGRQ-PVEAQOESCYGRLYSVKVNDDCNVEICQSVPEYATVGS 128
Db 61 KHTGAFCDKEFYCEGRQ-PVEAQOESCYGRLYSVKVNDDCNVEICQSVPEYATVGS 128
QY 129 PIETLATGKKDCVNVVITQOLPCEVEFVSSDPATPTSDSKLIWTIDCLGKCKITVW 188
Db 121 PIETLATGKKDCVNVVITQOLPCEVEFVSSDPATPTSDSKLIWTIDCLGKCKITVW 188
QY 189 VKPLKEGCCCTAATVACAPELRSYTKCGQPAICIKOEGPDCACLRCPVCYKIEVNTGSA 248
Db 181 VKPLKEGCCCTAATVACAPELRSYTKCGQPAICIKOEGPDCACLRCPVCYKIEVNTGSA 248

```

```
QY 249 IARNVTVDPVPGYSHASGQVLSFNLDMPGDKKVFTEFCQRRQITNATVTC 308
Db 241 IARNVVDPVPGYTHASGQVLSFNLDMPGDKSKCFEFCQRRKGTINATVTC 300
QY 309 GGHKCSANVTTVNPPCVQVNIAGDWSYVCKPVEYSISVSNPGDLVLHDVVQDTLP 368
Db 301 GGHKCSANVTTVNPPCVQVNIAGDWSYVCKPVEYTIWSNLGDLKLDYVVVDTVPSG 360
QY 369 VTVLEAPGECICCNKVVWRIKEMCPGETLQFKLVKVAQVPGRTNOVAVTSSENGCT 428
Db 361 ATILEAEGAEICCNKAVWCIREKCPGETLQFKLVKVAQSPGRTNOVAVTSSENGCT 420
QY 429 CAETTTHWKGAAATHMVCVLDNDPICVGTNTVYRICVTVNRGSAEDTNSLILKFSK 488
Db 421 CAETTTHWKGAAATHMVCVLDNDPICVGTNTVYRICVTVNRGSAEDTNSLILKFSK 480
QY 489 IASSGPTKGTISGNTVWFDPALPKLGSKEVSFVSVTLKGIAPGDARGEAILSSDT 542
Db 481 VSSGPTKGTITGNTVWFDPALPKLGSKEVSFVSVTLKGIAPGDARGEAILSSDT 534

RESULT 4
Q9ZAW0 PRELIMINARY: PRT; 534 AA.
AC Q9ZAW0:
DT 01-MAY-1999 (TremBLrel. 10, Created)
DT 01-MAY-1999 (TremBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE CYSTEIN-RICH OUTER MEMBRANE PROTEIN OMP-2 (FRAGMENT).
OS Chlamydomophila pecorum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OX NCBI_TaxID=85991;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W73;
RA Sheehy N., Markey B., Quinn P.J.;
RT "Sequence analysis of C. psittaci and C. pecorum 60 kDa genes.";
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U76761; AAD09598.1;
DR InterPro: IPR003506; Chlam_OMP6.
DR PRINTS: PR01336; CHLAMIDIAOM6.
FT NON_TER 1
FT NON_TER 534
SQ SEQUENCE 534 AA; 57298 MW; 39299DBD6A8534258 CRC64;
```

```
QY 10 TVLALTSMAFCASGGIEAAVAESLITKIVASAETKPAPVPTAKKVLVRRNKOPVE-- 67
Db 1 TVLALTSMTSSASGGIEAAVAESLITKIVANAESKRA---ADHKKEAKHNKKNKSKAT 57
QY 68 ---QKSRGAFCDKEFPCEEGRCQVPAEQESCGRLYSVKVNDDCNVETCSQVPEYATV 124
Db 58 KHSRNPSCCDKEF-PCEKNQQQVESRQESCGRMYSIKVNDDCNVELSQSVPEYATV 116
QY 125 GSPYPIEILAIAGKCDVVIITQQLPCEAEFVSSDPTTSDGKLWKIDRLGAGDKCK 184
Db 117 GSPYPIEILAIAGKCDVVIITQQLPCEAEFVSSDPTTADGKLWKIDRLGAGGERCK 176
QY 185 ITVWVKPLKEGCGCFTAAATVCAPCELSYTKCGOPALCIKQEGPDCACLRCPVCYKLEVN 244
Db 177 ITVWVKPIKEGCGCFTAAATVCAPCEIRSYTKCGOPALCIKQEGPDCACLRCPVCYKLEVN 236
QY 245 TGSIAIRNVTVDPVPGYSHASGQVLSFNLDMPGDKKVFTEFCQRRQITNATV 304
Db 237 TGSIAIRNVTVDPVPGYSHASGQVLSFNLDMPHGESKVLCEFCQRRQITNATV 296
QY 305 VTYCGGHKCSANVTTVNPPCVQVNIAGDWSYVCKPVEYSISVSNPGDLVLHDVV 364
Db 301 VTYCGGHKCSANVTTVNPPCVQVNIAGDWSYVCKPVEYISVSNPGDLVLHDVV 358
```

```
Db 297 VSYCGGHKCSANVTVINEPCVNVNVTISGVDWAYVCKPVEYITVSNPGDLVLRNVV 356
QY 365 LPSGVTVLEAPGECICCNKVVWRIKEMCPGETLQFKLVKVAQVPGRTNOVAVTSSENG 424
Db 357 LPSGALILEAAGAEICCNKAVWCIREKCPGETLQFKLVKVAQVPGOITISQVNVN 416
QY 425 TCTCAETTTTHWKGAAATHMVCVLDNDPICVGTNTVYRICVTVNRGSAEDTNSLILKFSK 484
Db 417 VCTSCADATTYWKGLAATHMVCVLDNDPICVGTNTVYRICVTVNRGSAEDTNSLILKFSK 476
QY 485 ELQPIASSGPTKGTISGNTVWFDPALPKLGSKEVSFVSVTLKGIAPGDARGEAILSSDT 542
Db 477 ELQPISSGPTKGTISGNTVWFDPALPKLGSKEVSFVSVTLKGIAPGDARGEAILSSDT 534

RESULT 5
Q9PJUV0 PRELIMINARY: PRT; 554 AA.
AC Q9PJUV0:
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE 60 KDA OUTER MEMBRANE PROTEIN.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MOPN / NIG;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RL Nucleic Acids Res. 28:1397-1406(2000).
DR EMBL: AE002341; AAF39537.1;
DR TIGR: TC0727;
KW Complete proteome.
SQ SEQUENCE 554 AA; 59784 MW; 1CF81B2471C27FD4 CRC64;
```

```
Query Match 73.7%; Score 2176; DB 16; Length 554;
Best Local Similarity 72.8%; Pred. No. 9.7e-170;
Matches 405; Conservative 74; Mismatches 69; Indels 8; Gaps 3;

QY 1 MSKLIRRVTVLALTSMAFCASGGIEAAVAESLITKIVASAETKPAPVPTAKKVLV 60
Db 7 MNKLIRRAVTIPAVTSVASLFAVSGVLETSMAESLSTNVISLADTK-AKETTSQKDRK 65
QY 61 RNKQPVQKSRGAFCDKEFPCEEGRCQVPAEQESCGRLYSVKVNDDCNVEICQSVPE 120
Db 66 KKHQ-----NRTSVVRKEVTAVRDTKA--VEPRQDSCFGKMYTVKVNDDNRVIVQSVPE 118
QY 121 YATVSGPYPIEILAIAGKCDVVIITQQLPCEAEFVSSDPTTSDGKLWKIDRLGAG 180
Db 119 YATVSGPYPIEILAIAGKCDVVIITQQLPCEAEFVSSDPTTADGKLWKIDRLGAG 178
QY 181 DKCKITVWVKPLKEGCGCFTAAATVCAPCELSYTKCGOPALCIKQEGPDCACLRCPVCYK 240
Db 179 ESKITVWVKPLKEGCGCFTAAATVCAPCEIRSVTKCGOPALCIKQEGPDCACLRCPVYRI 238
QY 241 EVVNTGSAIRNVTVDPVPGYSHASGQVLSFNLDMPGDKKVFTEFCQRRQIT 300
Db 239 NVVNOGTATARNVVDPVPGYAHASGQVLSYTLGDMQPGEOITITVTEFCPLKGRVT 298
QY 301 NVATVTVCGGHKCSANVTTVNPPCVQVNIAGDWSYVCKPVEYSISVSNPGDLVLHDVV 360
Db 299 NIATVTVCGGHKCSANVTTVNPPCVQVNIAGDWSYVCKPVEYISVSNPGDLVLHDVV 358
```

```
QY 361 IQDTLPSGVTVLEAPGGEICCNKVVRIKEMCPGETLOFLKLVKAQVPGPRTNOVATSE 420
Db 359 IEDTLSPGTVVBAAGAAISCNKLKLVMTLKLNPESLQYKLVVRAQTPGQFTNNVVKSC 418
QY 421 SNGCTCTCAETTHWKGLAATHMVCVLDNDPDCVGVNTYRICVTVNRGSAEDTVNSLIL 480
Db 419 SDGIGTCAEATYTWKGVAAATHMVCVVDTCDPICVGVNTYRICVTVNRGSAEDTVNSLIL 478
QY 481 KFSKELQPIASSGPTKGTISGNTVFDALPKLGSKEVSFVTLKGIAPGDARGEAI 540
Db 479 KFSKELQPISSGPTKGTITGNTVFDLPLRLGSKETVEFSVTLKAYSAGDARGEAILSS 538
QY 541 DLTLPSPVSDTENTHY 556
Db 539 DLTLPSPVSDTENTHIY 554

RESULT 6
Q93317 PRELIMINARY: PRT; 547 AA.
AC Q93317;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MAJOR OUTER MEMBRANE PROTEIN.
GN OMPA.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UW-12, AND UW-36;
RX MEDLINE=21450826; PubMed=11567000;
RA Millman K.B., Taware S., Dean D.;
RT "Recombination in the ompA Gene but Not the ompB Gene of Chlamydia
RT Contributes to Serovar-Specific Differences in Tissue Tropism, Immune
RT Surveillance, and Persistence of the Organism.";
RL J. Bacteriol. 183:5997-6008(2001).
DR EMBL; AF304330; AAL14100.1; -.
DR EMBL; AF304329; AAL14099.1; -.
SQ SEQUENCE 547 AA; 58626 MW; 817BA3000FEA6A71 CRC64;
```

```
Query Match 73.1%; Score 2157.5; DB 2; Length 547;
Best Local Similarity 71.9%; Pred. No. 3.1e-168;
Matches 402; Conservative 73; Mismatches 69; Indels 15; Gaps 4;

QY 1 MSLKIRRVTVLALTSMAFCASGAGGIEAAVAESLIITKIVASAEKPAVPMTAKKVLRLR 60
Db 1 MNKLIRRAVTIFAVTSVLSFASGLVETSMAESLSTNVISLADTKAK--DNTSHKSKAR 58

QY 61 RN--KQPVQKSRGAFCDKEFPCEEGRCQPVQAEQESCYGRYSVKVYNDNCNVEICQS 117
Db 59 KKHSETPVDR-----KEVAPVHESKA--TGPQKQDCFGMYTVKVNDRNVEITQA 108

QY 118 VPEYATVGSYPYIEILAIAGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRL 177
Db 109 VPEYATVGSYPYIEITATCKRDCVDVITQQLPCEAEFVRSDPATPTADGKLVWKIDRL 168

QY 178 GAGDKCKITVWKLKEGCGCCTAATVCACPELRSYTKGQPAICIKOEGPDCACLRCPVC 237
Db 169 GQGEKSKITVWKLKEGCGCCTAATVCACPELRSYTKGQPAICVQKQEPENACLRCPV 228

QY 238 YKIEVNTVGSATARNVTVNDPNPDGYSHASQORVLSFNLGDMRPGDKKVTVEFCQRRG 297
Db 229 YKINVNQGTARNVTVNDPNPDGYSHASQORVLSFNLGDMRPGDKKVTVEFCQRRG 288

QY 298 QITNATVTVYCGGHKSNANVTWVNEPCVQVNISGADWSYCKPVEYSISVSNPGLVLH 357
Db 289 CATNIATVSYCGGHKSNATVTVINEPCVQVSIAGADWSYCKPVEYSISVSNPGLVLR 348

QY 358 DVIQDTLPSGVTVLEAPGGETCCNKVVRKEMCPGETLOFLKLVKAQVPGPRTNOVAV 417
Db 357 DVIQDTLPSGVTVLEAPGGETCCNKVVRKEMCPGETLOFLKLVKAQVPGPRTNOVAV 417
```

```
Db 349 DVVVEDTLSPGVTVLEAAGAAQISCNKVVWTVKELNPESLQYKLVVRAQTPGQFTNNVVV 408
QY 418 TSNSCGTCTCAETTHWKGLAATHMVCVLDNDPDCVGVNTYRICVTVNRGSAEDTVNS 477
Db 409 KSCDGGTCTCAEATYTWKGVAAATHMVCVVDTCDPVCGVNTYRICVTVNRGSAEDTVNS 468
QY 478 LILKFSKELQPIASSGPTKGTISGNTVFDALPKLGSKEVSFVTLKGIAPGDARGEAI 537
Db 469 LMLKFSKELQPVFSGPTKGTITGNTVFDLPLRLGSKETVEFSVTLKAYSAGDARGEAI 528
QY 538 LSSDTLTSPVSDTENTHY 556
Db 529 LSSDTLTSPVSDTENTHIY 547

RESULT 7
Q9S6B3 PRELIMINARY: PRT; 395 AA.
AC Q9S6B3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE OUTER MEMBRANE PROTEIN 2 (FRAGMENT).
GN OMP2.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AR388;
RA Kaltenboeck B., Gao D.;
RT "PCR amplification and sequencing of the partial coding region of the
RT outer membrane protein 2 Omp2 of Chlamydia.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF111201; AAD20337.1; -.
FT NON_TER 1
FT NON_TER 395
FT NON_TER 395
SQ SEQUENCE 395 AA; 42272 MW; 7AD2B76F3F06EA37 CRC64;
```

```
Query Match 72.1%; Score 2127; DB 2; Length 395;
Best Local Similarity 100.0%; Pred. No. 6.5e-166;
Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 IGKKDCVDVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGDKCKITVWVKPLKE 194
Db 1 IGKKDCVDVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGDKCKITVWVKPLKE 60

QY 195 GCCFTAAATVCACPELRSYTKGQPAICIKOEGPDCACLRCPVCYKTEVNTGSAIARNVT 254
Db 61 GCCFTAAATVCACPELRSYTKGQPAICIKOEGPDCACLRCPVCYKTEVNTGSAIARNVT 120

QY 255 VDNVPDGYSHASQORVLSFNLGDMRPGDKKVTVEFCQRRGQITNATVTVYCGGHKCS 314
Db 121 VDNVPDGYSHASQORVLSFNLGDMRPGDKKVTVEFCQRRGQITNATVTVYCGGHKCS 180

QY 315 ANVTTVNEPCVQVNISGADWSYCKPVEYSISVSNPGLVLDVHVDIOTLPSGVTVLEA 374
Db 181 ANVTTVNEPCVQVNISGADWSYCKPVEYSISVSNPGLVLDVHVDIOTLPSGVTVLEA 240

QY 375 PGGETCCNKVVRKEMCPGETLOFLKLVKAQVPGPRTNOVATSESNCGTCTCAETTT 434
Db 241 PGGETCCNKVVRKEMCPGETLOFLKLVKAQVPGPRTNOVATSESNCGTCTCAETTT 300

QY 435 HWKGLAATHMVCVLDNDPDCVGVNTYRICVTVNRGSAEDTVNSLILKFSKELQPIASSGP 494
Db 301 HWKGLAATHMVCVLDNDPDCVGVNTYRICVTVNRGSAEDTVNSLILKFSKELQPIASSGP 360

QY 495 TKGTISGNTVTVFDALPKLGSKEVSFVTLKGIAP 529
Db 361 TKGTISGNTVTVFDALPKLGSKEVSFVTLKGIAP 395
```

```
RESULT 8
Q93FV8 PRELIMINARY; PRT; 511 AA.
ID Q93FV8;
AC Q93FV8;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE 60 KDA CYSTEINE-RICH OMP (FRAGMENT).
EN OMCB.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HAR-13;
RX MEDLINE=21450826; PubMed=11567000;
RA Millman K.L., Tavare S., Dean D.;
RT "Recombination in the ompA Gene but Not the ompB Gene of Chlamydia
RT Contributes to Serovar-Specific Differences in Tissue Tropism, Immune
RT Surveillance, and Persistence of the Organism.";
RL J. Bacteriol. 183:5997-6008(2001).
DR EMBL; AF304332; AAL14102.1; -.
FT NON_TER 511
SQ SEQUENCE 511 AA; 54978 MW; 27555CB67AA9214B CRC64;

Query Match 68.1%; Score 2008.5; DB 2; Length 511;
Best Local Similarity 71.1%; Pred. No. 4.4e-156;
Matches 372; Conservative 71; Mismatches 65; Indels 15; Gaps 4;

QY 1 MSKLIRRVTVLALTSMAFCFASGIEAIAVAESLITKIVASAEATKPAVPMTAKKVLVR 60
DB 1 MNKLIRRAVTVFAVTSVASFASGLVETSMASLSTNVISLADTKAK--DNTSHKSKAR 58
QY 61 RN---KOPVEOKSRGAFCKDFEYCEGRQCPVEAQAQESCYGRLYSVKVNDCNVEICQS 117
DB 59 KNSKETPVDR-----KEVAPVHESKA--TGPQKDSFCFRMYTVKVNDRNVEITQA 108
QY 118 VPEYATVGSPPYIEILAIGKDCVDVVTITQOLPCEAEFVSSDPETPTSDGKLWKIDRL 177
DB 109 VPEYATVGSPPYIEITATGKRDVDVITITQOLPCEAEFVSSDPATPTADGKLWKIDRL 168
QY 178 GAGDKCKITVWKPLKEGCCFTATVCACPELRSYTKCGQPAICIKQEGPCACLRCPVC 237
DB 169 GQGEKSKITVWKPLKEGCCFTATVCACPELRSYTKCGQPAICVQKQEGPENACLRCPV 228
QY 238 YKIEVNTGSAIARNTVNDVPDGYSHASQORVLSFNLGDMRPGDKKVFVEFCPQRRG 297
DB 229 YKINIVNGTATARNVVYENVPDGYAHSSGORVLTFTLGDMPQGEHRTITVEFCPLKRG 288
QY 298 QITNATVTCGGHKCSANVTTVNNEPCVQVNSIGADMSYVCKPVEYSTSVSNPGDLVLH 357
DB 289 RATNATVTCYCGGHKNTASVTTVINNEPCVQVNSIGADMSYVCKPVEYSTSVSNPGDLVL 348
QY 358 DVVIQDTLPSGVTVLEAPGGECCNKVVRKEMCPGETLQFKLVVKAQVPGGRFTNOVAV 417
DB 349 DVVVEDTLSPGVTVLEAGAOISCNKVVYKELNPGESLQFKLVRAQTPGQFTNNVVV 408
QY 418 TSESCGCTCAETTHWKGAAHMCVLDNDPICVGVNTVYICVTRNRSAGDNTNVS 477
DB 409 KSCSDCGTCSCAEATTVWKGAAHMCVVDTCDFPCVGVNTVYICVTRNRSAGDNTNVS 468
QY 478 LILKFSKELQPIASSGPTKGTISGNTVFPDALPKLGSKESVEF 520
DB 469 LMLKFSKELQPVFSFGPTKGTITGNTVFDLSPLRLGSKESVEF 511

RESULT 9
Q9X4I2 PRELIMINARY; PRT; 395 AA.
ID Q9X4I2;
AC Q9X4I2;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
```

```
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE OUTER MEMBRANE PROTEIN 2 (FRAGMENT).
GN OMP2.
OS Chlamydia abortus.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83555;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B577;
RA Kaltenboeck B., Gao D.;
RT "PCR amplification and sequencing of the partial coding region of the
RT outer membrane protein 2 Omp2 of Chlamydia.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF111200; AAD20336.1; -.
FT NON_TER 1
FT NON_TER 395
SQ SEQUENCE 395 AA; 42179 MW; FE4D01BB03B97F5 CRC64;

Query Match 65.0%; Score 1917; DB 2; Length 395;
Best Local Similarity 87.8%; Pred. No. 9.8e-149;
Matches 347; Conservative 24; Mismatches 24; Indels 0; Gaps 0;

QY 135 IGKKDCVDVVTITQOLPCEAEFVSSDPETPTSDGKLWKIDRLGAGDKCKITVWVKPLKE 194
DB 1 VGKKDCVNVVITQOLPCEVEFVSSDPATPTSDSKLIWTIDCLGQGEKCKITVWVKPLKE 60
QY 195 GCCCTAATVCAPELRSYTKCGQPAICIKQEGPCACLRCPVCYKIEVVNTGSAIARNVT 254
DB 61 GCCCTAATVCAPELRSYTKCGQPAICIKQEGPCACLRCPVCYKIEVCNTGSAIARNVV 120
QY 255 VDNVPDGYSHASQORVLSFNLGDMRPGDKKVFVEFCPQRRGOITNATVTCYCGHKCS 314
DB 121 VDNVPDGYTHASQORVLSFNLGDMRPGDKSCFSVEFCPQKRGKITNATVTCYCGHKCS 180
QY 315 ANVTTVNNEPCVQVNSIGADMSYVCKPVEYSTSVSNPGDLVLDVVDIOTLPSGVTVLEA 374
DB 181 ANVTTVNNEPCVQVNSIGADMSYVCKPVEYSTSVSNPGDLVLDVVDIOTLPSGVTVLEA 240
QY 375 PGGECCNKVVRKEMCPGETLQFKLVVKAQVPGGRFTNOVAVTSESCGCTCAETTT 434
DB 241 EGAEICCNKAVWCIKEMCPGETLQFKLVVKAQSGPKFTNQVVVKTNSDGTCTSCAEAT 300
QY 435 HWKGAAHMCVLDNDPICVGVNTVYICVTRNRSAGDNTNVSLLILKFSKELQPIASSGP 494
DB 301 HWKGAAHMCVLDNDPICVGVNTVYICVTRNRSAGDNTNVSLLILKFSKELQPVSSSGP 360
QY 495 TKGITSGNTVFPDALPKLGSKESVEFVTLKGIAPI 529
DB 361 TKGITGNTVFPDALPKLGSKESVEFVTLKGIAPI 395

RESULT 10
Q9X4I1 PRELIMINARY; PRT; 395 AA.
ID Q9X4I1;
AC Q9X4I1;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE OUTER MEMBRANE PROTEIN 2 (FRAGMENT).
GN OMP2.
OS Chlamydia pecorum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=85991;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FCSTRA;
RA Kaltenboeck B., Gao D.;
RT "PCR amplification and sequencing of the partial coding region of
RT homologs of the outer membrane protein 2 Omp2 of Chlamydia.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF111199; AAD20335.1; -.
FT NON_TER 1
FT NON_TER 1
```

```
FT NON_TER 395 395
SQ SEQUENCE 395 AA; 42205 MW; 6A90D7A41575A89E CRC64;

Query Match 62.1%; Score 1833; DB 2; Length 395;
Best Local Similarity 83.5%; Pred. No. 7.3e-142;
Matches 330; Conservative 34; Mismatches 31; Indels 0; Gaps 0;

QY 135 IGKKKDCVDVITQOLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGDKCKITVWVKPLKE 194
Db 1 IGKKKRCNVVITQOLPCEAEFVSSDPETPTADGKLMWKIDRLGQGERCKITVWVKPLKE 60

QY 195 GCCTAATVACAPERSYTKGCGPAICIKQEGPDACACLRCPVCYKIEVNTGSAIARNVT 254
Db 61 GCCTAATVACAPERSYTKGCGPAICIKQEGPDACACLRCPVCYKIEVNTGSAIARNV 120

QY 255 VDNVPDGYSHASQORVLSFNLGDMRPGDKKVFTEFCPPQRGGQITTNVATVTCYCGHKCS 314
Db 121 VDNVPDGYSHASQORVLSFNLGDMRPGDKKVFTEFCPPQRGGQITTNVATVTCYCGHKCS 180

QY 315 ANVTTVNPEPCVQVNIISGADMSYVCKPVEYSISVSNPGDLVLHDVVIQDTLPSPGVTVLEA 374
Db 181 ANVTTVNPEPCVQVNIISGADMSYVCKPVEYSISVSNPGDLVLHDVVIQDTLPSPGVTVLEA 240

QY 375 PGGEICCNKVVWRKEMCPGETLQFKLVKQAQVPGRETNQAVTSESNCGTCTCAETTT 434
Db 241 AGAEISCNKAVKWKELCPGETLQFKLVKQAQVPGRETNQAVTSESNCGTCTCAETTT 300

QY 435 HWKGLAATHMCVLDNDPICVGNTRYICVTNRGSAEDTNVSLILKPSKELQPIASSGP 494
Db 301 YWKGLAATHMCVLDNDPICVGNTRYICVTNRGSAEDTNVSLILKPSKELQPIASSGP 360

QY 495 TKGTISGNTVFDPALPKLGSKESEFVSFVTLKGIAP 529
Db 361 TKGTISGNTVFDPALPKLGSKESEFVSFVTLKGVAP 395

RESULT 11
Q46164 ID Q46164 PRELIMINARY; PRT; 183 AA.
AC Q46164;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE OMPB (FRAGMENT).
GN OMPB.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83358;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1.
RA Hotzel H., Grossmann E., Mutschmann F., Sachse K.;
RT "Genetic characterization of a Chlamydia pneumoniae isolate from
an African frog and comparison to currently accepted biovars.";
RL Syst. Appl. Microbiol. 0:0-0(2001).
DR EMBL; AF347609; AAK30131.1; -.
DR InterPro; IPR003506; Chlam_OMP6.
DR PRINTS; PR01336; CHLAMIDIAOM6.
FT NON_TER 1
FT NON_TER 176
SQ SEQUENCE 176 AA; 19094 MW; EE48731EFD2A991E CRC64;

Query Match 31.2%; Score 920; DB 2; Length 176;
Best Local Similarity 98.9%; Pred. No. 1.3e-67;
Matches 174; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 13 ALTNASCFASGGIEAAVAESLITKIVASAEKTPAPVPTAKKVLRRNKQPVQKSRG 72
Db 1 ALTNASCFASGGIEAAVAESLITKIVASAEKTPAPVPTAKKVLRRNKQPVQKSRG 60

QY 73 AFCDKEFYPCBEGRCQPVQAQESCYGRLYSVKVNDDCNVCEICQSVPEATVGSPIEI 132
Db 61 AFCDKEFYPCBEGRCQPVQAQESCYGRLYSVKVNDDCNVCEICQSVPEATVGSPIEI 120

QY 133 LAIGKKDCVDVITQOLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGDKCKITVW 188
Db 121 LAIGKKDCVDVITQOLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGDKCKITVW 176

RESULT 13
Q9AF83 ID Q9AF83 PRELIMINARY; PRT; 182 AA.
AC Q9AF83;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE OUTER MEMBRANE PROTEIN 2 (FRAGMENT).
GN OMP2.
OS Chlamydia felis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83356;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FP;
RX MEDLINE=21417404; PubMed=11526131;
RA Hartley J.C., Kaye S., Stevenson S., Bennett J., Ridgway G.;
RT "Per detection and molecular identification of Chlamydiaceae
species.";
RT
```

```

J. Clin. Microbiol. 39:3072-3079(2001).
EMBL: AF367407; AAK38115.1; -.
DR InterPro: IPR003506; Chlam_OMP6.
DR PRINTS: PRO1336; CHLAMIDIAOM6.
FT NON_TER 1
FT NON_TER 182
FT NON_TER 182
SQ SEQUENCE 182 AA; 19903 MW; 9CEB2D5192EBFICA CRC64;

Query Match      23.2%; Score 685; DB 2; Length 182;
Best Local Similarity 71.4%; Pred. No. 2.3e-48;
Matches 130; Conservative 19; Mismatches 31; Indels 2; Gaps 2;

QY 12 LALTSNASCFAASGGIEAAVAESLITKIVASAEKTPAPV-PMTAKKVLVRNKPVEQKS 70
DB 1 LALTSNASCFAASGGIEAAVAESLITKIVASAEKTPAPV-PMTAKKVLVRNKPVEQKS 70
QY 71 RGAFCDKEFPCEGRQCO-PVEAQOESCYGRLYSVKYNDDCNVEICQSVPEYATVGSYP 129
DB 1 RGAFCDKEFPCEGRQCO-PVEAQOESCYGRLYSVKYNDDCNVEICQSVPEYATVGSYP 129
QY 61 NNAFCDKFEPCEGGQCSVDTRQESYCGKMSYRVNDDCNVEISQAVPEYATVGSYP 120
DB 1 NNAFCDKFEPCEGGQCSVDTRQESYCGKMSYRVNDDCNVEISQAVPEYATVGSYP 120
QY 130 IEILAIGKKDCVDVVIITQQLPCEAEFVSSDPETPTSDGKLVKIDRLGAGDKCKITVW 189
DB 1 IEILAIGKKDCVDVVIITQQLPCEAEFVSSDPETPTSDGKLVKIDRLGAGDKCKITVW 189
QY 190 KP 191
DB 181 KP 182

RESULT 14
QYRB53 PRELIMINARY; PRT; 128 AA.
AC QYRB53;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE OUTER MEMBRANE PROTEIN B (FRAGMENT).
GN OMPB.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=833558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FROG-MI-L; PubMed=10364623;
RA Berger L., Volp K., Mathews S., Speare R., Timms P.;
RT "Chlamydia pneumoniae in a free-ranging giant barred frog (Mixophyes
iteratus) from Australia.";
RL J. Clin. Microbiol. 37:2378-2380(1999).
DR EMBL: AF102831; AAD25999.1; -.
FT NON_TER 1
FT NON_TER 128
FT NON_TER 128
SQ SEQUENCE 128 AA; 14112 MW; 0C38639151098A70 CRC64;

Query Match      22.8%; Score 674; DB 2; Length 128;
Best Local Similarity 98.4%; Pred. No. 1.2e-47;
Matches 126; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 38 IVASAEKTPAPVMTAKKVLVRNKPVEQKSRGAFCDKEFPCEGRQCPVEAQOESC 97
DB 1 IVASAEKTPAPVMTAKKVLVRNKPVEQKSRGAFCDKEFPCEGRQCPVEAQOESC 97
QY 98 YGRLYSVKYNDDCNVEICQSVPEYATVGSYPPIEILAIGKKDCVDVVIITQQLPCEAEFVS 137
DB 1 YGRLYSVKYNDDCNVEICQSVPEYATVGSYPPIEILAIGKKDCVDVVIITQQLPCEAEFVS 137
QY 158 SDPETPT 165
DB 121 SDPETPT 128
```

```

RESULT 15
Q46163 PRELIMINARY; PRT; 178 AA.
ID Q46163;
AC Q46163;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE OMPB (FRAGMENT).
GN OMPB.
OS Chlamydia pecorum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=85991;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K13, MUT13, MUT133, GC522, AND GC232;
RA Glassick T., Giffard P.M., Timms P.M.;
RT "Outer membrane protein 2 gene sequences indicate that Chlamydia
pecorum and Chlamydia pneumoniae cause infections in koalas.";
RL Syst. Appl. Microbiol. 0:0-0(1996).
DR EMBL: U56924; AAB00571.1; -.
FT NON_TER 1
FT NON_TER 178
FT NON_TER 178
SQ SEQUENCE 178 AA; 19490 MW; FF57E90C9D58B691 CRC64;

Query Match      22.7%; Score 669; DB 2; Length 178;
Best Local Similarity 72.2%; Pred. No. 4.6e-47;
Matches 130; Conservative 21; Mismatches 21; Indels 8; Gaps 2;

QY 13 ALTSNASCFAASGGIEAAVAESLITKIVASAEKTPAPVMTAKKVLVRNKPVE----- 67
DB 1 ALTSNASCFAASGGIEAAVAESLITKIVASAEKTPAPVMTAKKVLVRNKPVE----- 67
QY 68 OKSRGAFCDKEFPCEGRQCPVEAQOESCYGRLYSVKYNDDCNVEICQSVPEYATVGS 127
DB 1 OKSRGAFCDKEFPCEGRQCPVEAQOESCYGRLYSVKYNDDCNVEICQSVPEYATVGS 127
QY 58 SRNFCSCDKEFPCEGRQCPVEAQOESCYGRLYSVKYNDDCNVEICQSVPEYATVGS 117
DB 1 SRNFCSCDKEFPCEGRQCPVEAQOESCYGRLYSVKYNDDCNVEICQSVPEYATVGS 117
QY 128 YPIETILAIGKKDCVDVVIITQQLPCEAEFVSSDPETPTSDGKLVKIDRLGAGDKCKITV 187
DB 1 YPIETILAIGKKDCVDVVIITQQLPCEAEFVSSDPETPTSDGKLVKIDRLGAGDKCKITV 187
QY 118 YPIETILAIGKKDCVDVVIITQQLPCEAEFVSSDPETPTSDGKLVKIDRLGAGDKCKITV 177
DB 1 YPIETILAIGKKDCVDVVIITQQLPCEAEFVSSDPETPTSDGKLVKIDRLGAGDKCKITV 177

Search completed: May 25, 2002, 22:23:59
Job time: 502 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2002, 22:26:03 ; Search time 31.45 Seconds
(without alignments)
3058.352 Million cell updates/sec

Title: US-09-523-647-2
Perfect score: 556
Sequence: 1 MSKLRRVVTVLALTSAMSC.....ILSSDFTLTSPVSDTENTHYV 556

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 562222 seqs, 172994929 residues

Word size : 12
Total number of hits satisfying chosen parameters: 19

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL19.*

1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	395	71.0	395	2 Q9S6B3	Q9S6B3 chlamydia p
2	134	24.1	183	2 Q46164	Q46164 chlamydia p
3	127	22.8	176	2 Q9AG10	Q9AG10 chlamydia p
4	104	18.7	128	2 Q9RB53	Q9RB53 chlamydia p
5	46	8.3	395	2 Q9X412	Q9X412 chlamydia p
6	46	8.3	534	2 Q9ZAW1	Q9ZAW1 chlamydia p
7	46	8.3	557	2 Q9AIS7	Q9AIS7 chlamydia p
8	45	8.1	558	2 P94664	P94664 chlamydia m
9	39	7.0	554	16 Q9PJV0	Q9PJV0 chlamydia m
10	31	5.6	534	2 Q9ZAW0	Q9ZAW0 chlamydia p
11	26	4.7	178	2 Q46163	Q46163 chlamydia p
12	26	4.7	178	2 Q46165	Q46165 chlamydia p
13	25	4.5	511	2 Q93FV8	Q93FV8 chlamydia t
14	25	4.5	547	2 Q933I7	Q933I7 chlamydia t
15	24	4.3	395	2 Q9X411	Q9X411 chlamydia p
16	17	3.1	157	2 Q9AF85	Q9AF85 chlamydia p

17 17 3.1 157 2 Q9AF84 chlamydia p
18 17 3.1 182 2 Q9AF83 chlamydia p
19 16 2.9 173 2 Q9AF82 chlamydia s

ALIGNMENTS

RESULT 1
Q9S6B3 PRELIMINARY; PRT; 395 AA.
ID Q9S6B3;
AC Q9S6B3;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE OUTER MEMBRANE PROTEIN 2 (FRAGMENT).
GN OMP2.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AR388;
RA Kaltenboeck B., Gao D.;
RT "PCR amplification and sequencing of the partial coding region of the outer membrane protein 2 Omp2 of Chlamydia";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF111201; AAD20337.1; -.
FT NON_TER 1
FT NON_TER 395
SQ SEQUENCE 395 AA; 42272 MW; 7AD2B76F3F06EA37 CRC64;

Query Match 71.0%; Score 395; DB 2; Length 395;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 IGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLWVKIDRLGAGDKCKITVWVKPLKE 194
DB 1 IGKDCVDVITQQLPCEAEFVSSDPETPTSDGKLWVKIDRLGAGDKCKITVWVKPLKE 60
QY 195 GCFTAAATVCACPELRSYTKGQPAICIKQEGDCACLRCPVCYKIEVVTGSAIARNVT 254
DB 61 GCFTAAATVCACPELRSYTKGQPAICIKQEGDCACLRCPVCYKIEVVTGSAIARNVT 120
QY 255 VDNVPDGYSHASGORVLSFNLGDMRPGDKVFTVEFCPQRRGQITNATVTCGGHKCS 314
DB 121 VDNVPDGYSHASGORVLSFNLGDMRPGDKVFTVEFCPQRRGQITNATVTCGGHKCS 180
QY 315 ANVTTVNEPCVQVNLISGADWSYCKPVEYSISVSNPGDLVLDVVIQDTLPFGVTVLEA 374
DB 181 ANVTTVNEPCVQVNLISGADWSYCKPVEYSISVSNPGDLVLDVVIQDTLPFGVTVLEA 240
QY 375 PGGEICCNKVVWRIKEMCPGETIQFKLVVKAQVPGRFTNQAVTSESNCGTCTSCAETTT 434
DB 241 PGGEICCNKVVWRIKEMCPGETIQFKLVVKAQVPGRFTNQAVTSESNCGTCTSCAETTT 300
QY 435 HWKGLAATHMCLVDNDPCTCGENTVYRCVTVNRGSAEDTNVSLILKFSKLOPIASSGP 494
DB 301 HWKGLAATHMCLVDNDPCTCGENTVYRCVTVNRGSAEDTNVSLILKFSKLOPIASSGP 360
QY 495 TKGTSIGNTVVFDALPKLGSKESEVFSVTLKGIAP 529
DB 361 TKGTSIGNTVVFDALPKLGSKESEVFSVTLKGIAP 395
RESULT 2
Q46164 PRELIMINARY; PRT; 183 AA.
ID Q46164
AC Q46164;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

```

DE OMPB (FRAGMENT).
GN OMPB.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KC, AND 736;
RA Glassick T., Giffard P.M., Timms P.M.;
RT "Outer membrane protein 2 gene sequences indicate that Chlamydia
   pecorum and Chlamydia pneumoniae cause infections in koalas.";
RL Syst. Appl. Microbiol. 0:0-0(1996).
DR EMBL; U56925; AAB00572.1; -.
FT NON_TER 1
FT NON_TER 183
SQ SEQUENCE 183 AA; 19846 MW; 09B8D084F4C016E CRC64;

Query Match 24.1%; Score 134; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 6.9e-133;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOESC YGRLYSVKVNDDCNVEICQSVPEY 121
DB 50 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOESC YGRLYSVKVNDDCNVEICQSVPEY 109

QY 122 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGD 181
DB 110 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGD 169

QY 182 KCKITVWVPLKEG 195
DB 170 KCKITVWVPLKEG 183

RESULT 3
QY 62 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOESC YGRLYSVKVNDDCNVEICQSVPEY 121
DB 50 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOESC YGRLYSVKVNDDCNVEICQSVPEY 109
QY 122 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGD 181
DB 110 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGD 169
QY 182 KCKITVWVPLKEG 195
DB 170 KCKITVWVPLKEG 183

Query Match 22.8%; Score 127; DB 2; Length 176;
Best Local Similarity 100.0%; Pred. No. 1.7e-125;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOESC YGRLYSVKVNDDCNVEICQSVPEY 121
DB 50 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOESC YGRLYSVKVNDDCNVEICQSVPEY 109

QY 122 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGD 181
DB 110 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGD 169

```

```

QY 182 KCKITVW 188
DB 170 KCKITVW 176

RESULT 4
QY 62 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOESC YGRLYSVKVNDDCNVEICQSVPEY 121
DB 50 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOESC YGRLYSVKVNDDCNVEICQSVPEY 109
QY 122 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGD 181
DB 110 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGD 169
QY 182 KCKITVWVPLKEG 195
DB 170 KCKITVWVPLKEG 183

Query Match 18.7%; Score 104; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 2.5e-101;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOESC YGRLYSVKVNDDCNVEICQSVPEY 121
DB 50 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOESC YGRLYSVKVNDDCNVEICQSVPEY 84
QY 122 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGD 165
DB 85 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGD 128

RESULT 5
QY 62 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOESC YGRLYSVKVNDDCNVEICQSVPEY 121
DB 50 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOESC YGRLYSVKVNDDCNVEICQSVPEY 84
QY 122 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGD 165
DB 85 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGD 128

Query Match 8.3%; Score 46; DB 2; Length 395;
Best Local Similarity 100.0%; Pred. No. 1.3e-39;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOESC YGRLYSVKVNDDCNVEICQSVPEY 121
DB 50 NKQPVQKSRGAFCDKEFFPCBEGRCQPVEAQOESC YGRLYSVKVNDDCNVEICQSVPEY 109
QY 122 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGD 181
DB 110 ATVGSPYPIEIIAIGKKDCVDVVITQQLPCEAEFVSSDPETPTSDGKLVWKIDRLGAGD 169
QY 182 KCKITVWVPLKEG 195
DB 170 KCKITVWVPLKEG 183

Query Match 8.3%; Score 46; DB 2; Length 395;
Best Local Similarity 100.0%; Pred. No. 1.3e-39;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 182 KCKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKQEGP 227
|||||
Db 48 KCKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKQEGP 93

RESULT 6
Q9ZAW1
ID Q9ZAW1 PRELIMINARY; PRT; 534 AA.
AC Q9ZAW1;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CYSTEIN-RICH OUTER MEMBRANE PROTEIN OMP-2 (FRAGMENT).
OS Chlamydomophila abortus.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OX NCBI_TaxID=83555;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S26/3;
RA Sheehy N., Markey B., Quinn P.J.;
RT "Sequence analysis of C. psittaci and C. pecorum 60 kDa genes.";
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U76760; AAD09597.1; -.
DR InterPro; IPR003506; Chlam_OMP6.
DR PRINTS; PR01336; CHLAMIDIOM6.
FT NON_TER 1
FT NON_TER 534
SQ SEQUENCE 534 AA; 57137 MW; 27898208041E92BE CRC64;

Query Match 8.3%; Score 46; DB 2; Length 534;
Best Local Similarity 100.0%; Pred. No. 1.7e-39;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 182 KCKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKQEGP 227
|||||
Db 174 KCKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKQEGP 219

RESULT 7
Q9AIS7
ID Q9AIS7 PRELIMINARY; PRT; 557 AA.
AC Q9AIS7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE 60 KDA CYSTEINE-RICH MEMBRANE COMPLEX PROTEIN.
GN CMCB.
OS Chlamydomophila abortus.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OX NCBI_TaxID=83555;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EBA;
RX MEDLINE=21078680; PubMed=11211261;
RA Bush R.M., Everett K.D.;
RT "Molecular evolution of the Chlamydiaceae.";
RL Int. J. Syst. Evol. Microbiol. 51:203-220(2001).
DR EMBL; AF240773; AAG60550.1; -.
DR InterPro; IPR003506; Chlam_OMP6.
DR PRINTS; PR01336; CHLAMIDIOM6.
SQ SEQUENCE 557 AA; 59761 MW; 817A20B071AD9609 CRC64;

Query Match 8.3%; Score 46; DB 2; Length 557;
Best Local Similarity 100.0%; Pred. No. 1.8e-39;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 182 KCKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKQEGP 227
|||||
Db 183 KCKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKQEGP 228

RESULT 8
P94664
ID P94664 PRELIMINARY; PRT; 558 AA.
AC P94664;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE OUTER MEMBRANE PROTEIN 2.
GN OMP2.
OS Chlamydomophila caviae.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OX NCBI_TaxID=83557;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GUINEA PIG INCLUSION CONJUNCTIVITIS (GPIC);
RX MEDLINE=97075924; PubMed=8918247;
RA Hsia R.C., Bavoil P.M.;
RT "Sequence analysis of the omp2 region of Chlamydia psittaci strain
GPIC: structural and functional implications.";
RL Gene 176:155-162(1996).
DR EMBL; U41759; AAB41143.1; -.
DR InterPro; IPR003506; Chlam_OMP6.
DR PRINTS; PR01336; CHLAMIDIOM6.
SQ SEQUENCE 558 AA; 60213 MW; 452E2BD79078935C CRC64;

Query Match 8.1%; Score 45; DB 2; Length 558;
Best Local Similarity 100.0%; Pred. No. 2e-38;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 CKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKQEGP 227
|||||
Db 185 CKITVWVKPLKEGCGCFTAAATVCACPELRSYTKCGQPAICIKQEGP 229

RESULT 9
Q9PJVO
ID Q9PJVO PRELIMINARY; PRT; 554 AA.
AC Q9PJVO;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 60 KDA OUTER MEMBRANE PROTEIN.
GN TC0727.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MOPN / NIGG;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
pneumoniae AK39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
DR EMBL; AE002341; AAF39537.1; -.
DR TIGR; TC0727; -.
KW Complete proteome.
SQ SEQUENCE 554 AA; 59784 MW; 1CF81B2471C27FD4 CRC64;

Query Match 7.0%; Score 39; DB 16; Length 554;
Best Local Similarity 100.0%; Pred. No. 4.4e-32;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 451 DPICVGVTVYRICVTVNRGSAEDTNVSLILKFSKELOPI 489
|||||
Db 449 DPICVGVTVYRICVTVNRGSAEDTNVSLILKFSKELOPI 487

```

RESULT 10
Q9ZAW0 ID Q9ZAW0 PRELIMINARY; PRT; 534 AA.
AC Q9ZAW0:
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CYSTEIN-RICH OUTER MEMBRANE PROTEIN OMP-2 (FRAGMENT).
OS Chlamydomophila pecorum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OC NCBI_TaxID=85991;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VR268;
RA Glassick T., Giffard P.M., Timms P.M.;
RT "Outer membrane protein 2 gene sequences indicate that Chlamydia pecorum and Chlamydia pneumoniae cause infections in koalas.";
RL EMBL; U76761; AAD09598.1; -.
DR Syst. Appl. Microbiol. 0:0-0(1996).
DR PRINTS; PR01336; CHLAMIDIAOM6.
FT NON_TER 1 534
FT NON_TER 534 534
SQ SEQUENCE 534 AA; 57298 MW; 3929DBD6A8534258 CRC64;

Query Match 5.6%; Score 31; DB 2; Length 534;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 506 FDALPKLGSKESVFSTLKGTPAGDARGE 536
Db 498 FDALPKLGSKESVFSTLKGTPAGDARGE 528

RESULT 11
Q46163 ID Q46163 PRELIMINARY; PRT; 178 AA.
AC Q46163:
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE OMPB (FRAGMENT).
GN OMPB.
OS Chlamydomophila pecorum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OC NCBI_TaxID=85991;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K13, MUT13, MUT133, GC522, AND GC232;
RA Glassick T., Giffard P.M., Timms P.M.;
RT "Outer membrane protein 2 gene sequences indicate that Chlamydia pecorum and Chlamydia pneumoniae cause infections in koalas.";
RL Syst. Appl. Microbiol. 0:0-0(1996).
DR EMBL; U56924; AAB00571.1; -.
FT NON_TER 1 178
FT NON_TER 178 178
SQ SEQUENCE 178 AA; 19490 MW; FF57E90C9D58B691 CRC64;

Query Match 4.7%; Score 26; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. No. 8.8e-19;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 116 QSVPEYATVGSYPPIEILAIGKKDCV 141
Db 106 QSVPEYATVGSYPPIEILAIGKKDCV 131

RESULT 12
Q46165 ID Q46165 PRELIMINARY; PRT; 178 AA.

```

```

Q46165;
AC 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE OMPB (FRAGMENT).
GN OMPB.
OS Chlamydomophila pecorum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OC NCBI_TaxID=85991;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VR268;
RA Glassick T., Giffard P.M., Timms P.M.;
RT "Outer membrane protein 2 gene sequences indicate that Chlamydia pecorum and Chlamydia pneumoniae cause infections in koalas.";
RL Syst. Appl. Microbiol. 0:0-0(1996).
DR EMBL; U56927; AAB00574.1; -.
FT NON_TER 1 178
FT NON_TER 178 178
SQ SEQUENCE 178 AA; 19506 MW; E296350C9D58B691 CRC64;

Query Match 4.7%; Score 26; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. No. 8.8e-19;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 116 QSVPEYATVGSYPPIEILAIGKKDCV 141
Db 106 QSVPEYATVGSYPPIEILAIGKKDCV 131

RESULT 13
Q93FV8 ID Q93FV8 PRELIMINARY; PRT; 511 AA.
AC Q93FV8:
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 60 KDA CYSTEINE-RICH OMP (FRAGMENT).
GN OMCB.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OC NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HAR-13;
RX MEDLINE=21450826; PubMed=11567000;
RA Millman K.L., Tavares S., Dean D.;
RT "Recombination in the ompA gene but not the ompB gene of Chlamydia contributes to serovar-specific differences in tissue tropism, immune surveillance, and persistence of the organism.";
RL J. Bacteriol. 183:5997-6008(2001).
DR EMBL; AF304332; AAL14102.1; -.
FT NON_TER 511 511
FT NON_TER 511 511
SQ SEQUENCE 511 AA; 54978 MW; 27555CB67AA9214B CRC64;

Query Match 4.5%; Score 25; DB 2; Length 511;
Best Local Similarity 100.0%; Pred. No. 2.6e-17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 184 KITVWVKPLKGGCCFTAAATVCACPE 208
Db 175 KITVWVKPLKGGCCFTAAATVCACPE 199

RESULT 14
Q93317 ID Q93317 PRELIMINARY; PRT; 547 AA.
AC Q93317:
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

```

Search completed: May 25, 2002, 22:29:44
Job time: 221 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: May 25, 2002, 22:24:57 ; Search time 118.77 Seconds
(without alignments)
1647.733 Million cell updates/sec

Title: US-09-523-647-2
Perfect score: 556
Sequence: 1 MSKLIRRVTVLALTSMA SCFASGGIEAAV A E S L T K I V A S A E T K P A P V P M T A K K V R L Y R 60

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 3502263 seqs, 351980561 residues

Word size : 12

Total number of hits satisfying chosen parameters: 67

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending Patents_AA_Main:*

- 1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
- 22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
- 23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
- 24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
- 25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
- 26: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match	Length DB ID	Description	
1	556	100.0	556 15	US-09-198-452A-596	Sequence 596, App
2	556	100.0	556 19	US-09-523-647-2	Sequence 2, Appli
3	556	100.0	556 22	US-09-841-132-398	Sequence 398, App
4	556	100.0	559 18	US-09-438-185-559	Sequence 559, App
5	556	100.0	559 18	US-09-438-185A-559	Sequence 559, App
6	25	4.5	25 24	US-10-020-269-38	Sequence 38, Appl
7	25	4.5	553 22	US-09-841-132-441	Sequence 441, App

8	25	4.5	557 16	US-09-201-228A-1036	Sequence 1036, Ap
9	20	3.6	20 24	US-10-020-269-41	Sequence 41, Appl
10	19	3.4	19 24	US-10-020-269-39	Sequence 39, Appl
11	16	2.9	16 18	US-09-454-684-253	Sequence 253, App
12	16	2.9	16 18	US-09-454-684A-253	Sequence 253, App
13	16	2.9	16 18	US-09-461-705-7	Sequence 7, Appli
14	16	2.9	16 19	US-09-556-877-253	Sequence 253, App
15	16	2.9	16 19	US-09-598-419-253	Sequence 253, App
16	16	2.9	16 20	US-09-620-412A-253	Sequence 253, App
17	16	2.9	16 22	US-09-841-132-253	Sequence 43, Appl
18	16	2.9	16 24	US-10-020-269-43	Sequence 250, App
19	16	2.9	20 18	US-09-454-684-250	Sequence 250, App
20	16	2.9	20 18	US-09-454-684A-250	Sequence 250, App
21	16	2.9	20 19	US-09-556-877-250	Sequence 250, App
22	16	2.9	20 19	US-09-598-419-250	Sequence 250, App
23	16	2.9	20 20	US-09-620-412A-250	Sequence 250, App
24	16	2.9	20 22	US-09-841-132-250	Sequence 250, App
25	13	2.3	19 24	US-10-020-269-40	Sequence 40, Appl
26	12	2.2	12 18	US-09-454-684-252	Sequence 252, App
27	12	2.2	12 18	US-09-454-684A-252	Sequence 252, App
28	12	2.2	12 19	US-09-556-877-252	Sequence 252, App
29	12	2.2	12 19	US-09-598-419-252	Sequence 252, App
30	12	2.2	12 20	US-09-620-412A-252	Sequence 252, App
31	12	2.2	12 22	US-09-841-132-252	Sequence 251, App
32	12	2.2	16 18	US-09-454-684-251	Sequence 251, App
33	12	2.2	16 18	US-09-454-684A-251	Sequence 251, App
34	12	2.2	16 19	US-09-556-877-251	Sequence 251, App
35	12	2.2	16 19	US-09-598-419-251	Sequence 251, App
36	12	2.2	16 20	US-09-620-412A-251	Sequence 251, App
37	12	2.2	16 22	US-09-841-132-251	Sequence 241, App
38	12	2.2	20 18	US-09-454-684-241	Sequence 241, App
39	12	2.2	20 18	US-09-454-684-242	Sequence 242, App
40	12	2.2	20 18	US-09-454-684-246	Sequence 246, App
41	12	2.2	20 18	US-09-454-684-247	Sequence 247, App
42	12	2.2	20 18	US-09-454-684-249	Sequence 249, App
43	12	2.2	20 18	US-09-454-684A-241	Sequence 241, App
44	12	2.2	20 18	US-09-454-684A-242	Sequence 242, App
45	12	2.2	20 18	US-09-454-684A-246	Sequence 246, App

ALIGNMENTS

RESULT 1
US-09-198-452A-596
: Sequence 596, Application US/09198452A
: GENERAL INFORMATION:
: APPLICANT: Griflais, R.
: TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragme
: TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, pr
: TITLE OF INVENTION: and treatment of infection
: FILE REFERENCE: 9710-003-999
: CURRENT APPLICATION NUMBER: US/09/198.452A
: CURRENT FILING DATE: 1998-11-24
: NUMBER OF SEQ ID NOS: 6849
: SEQ ID NO 596
: LENGTH: 556
: TYPE: PRT
: ORGANISM: Chlamydia pneumoniae
US-09-198-452A-596

Query Match 100.0%; Score 556; DB 15; Length 556;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MSKLIRRVTVLALTSMA SCFASGGIEAAV A E S L T K I V A S A E T K P A P V P M T A K K V R L Y R 60	
Db	1	MSKLIRRVTVLALTSMA SCFASGGIEAAV A E S L T K I V A S A E T K P A P V P M T A K K V R L Y R 60	
Qy	61	RNKQPEOKSRGAFCDKEFYPCBGRQCPVEAQEQSCYGRLYSVKVNDDCNVEICQSVPE 120	
Db	61	RNKQPEOKSRGAFCDKEFYPCBGRQCPVEAQEQSCYGRLYSVKVNDDCNVEICQSVPE 120	

Db 241 EVVNTGSAIARNVTVDPNPVPGDGYSHASQORVLSFNLGDMRPGDKKVFVTEFCPQRRGQIT 300
QY 301 NVATVTCGGHKCSANVTTVVNEPCQVQVNI SGADWSYVCKPVEYSISVSNPGLVLDVV 360
Db 301 NVATVTCGGHKCSANVTTVVNEPCQVQVNI SGADWSYVCKPVEYSISVSNPGLVLDVV 360
QY 361 IODTLPSTVLEAPGGEICCNKVVWRIKEMCPGETLQFKLVVKAQVPGRFTNOVAVTSE 420
Db 361 IODTLPSTVLEAPGGEICCNKVVWRIKEMCPGETLQFKLVVKAQVPGRFTNOVAVTSE 420
QY 421 SNCGTCTCAEATTHWKGLAATHMVCVLDNDPICVGENTVYRICVTVNRGSAEDTNVSLIL 480
Db 421 SNCGTCTCAEATTHWKGLAATHMVCVLDNDPICVGENTVYRICVTVNRGSAEDTNVSLIL 480
QY 481 KFSKELOPIASSGPTKGTISGNTVVPDALPKLGSKESEVFSVTLKGIAPGDARGEAILSS 540
Db 481 KFSKELOPIASSGPTKGTISGNTVVPDALPKLGSKESEVFSVTLKGIAPGDARGEAILSS 540
QY 541 DTLTSPVSDTENTHY 556
Db 541 DTLTSPVSDTENTHY 556

RESULT 4

US-09-438-185-559
; Sequence 559, Application US/09438185
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185
; PRIOR FILING DATE: 1999-11-11
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 559
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-438-185-559

Query Match 100.0%; Score 556; DB 18; Length 559;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSLKIRRVTVLTALTSMSACFASGGTEAAVAESLITKIVASAETKPAVPMTAKKRVLR 60
Db 4 MSLKIRRVTVLTALTSMSACFASGGTEAAVAESLITKIVASAETKPAVPMTAKKRVLR 63
QY 61 RNKQPVQKSRGAFCDKEFYPCPEGRCQPVQAQOESCYGRLYSVKVNDCNVEICQSVPE 120
Db 64 RNKQPVQKSRGAFCDKEFYPCPEGRCQPVQAQOESCYGRLYSVKVNDCNVEICQSVPE 123
QY 121 YATVGSPIPIETILAIAGKDCVDVVIITQQLPCEAEFVSSDPETPTSDGKLWVKIDRLGAG 180
Db 124 YATVGSPIPIETILAIAGKDCVDVVIITQQLPCEAEFVSSDPETPTSDGKLWVKIDRLGAG 183
QY 181 DKCKITWWKPLKEGCGCFTAAATVCACPELRSYTKCGOPATCICKQEGPDCACLRCPVCYKI 240
Db 184 DKCKITWWKPLKEGCGCFTAAATVCACPELRSYTKCGOPATCICKQEGPDCACLRCPVCYKI 243
QY 241 EVVNTGSAIARNVTVDPNPVPGDGYSHASQORVLSFNLGDMRPGDKKVFVTEFCPQRRGQIT 300
Db 244 EVVNTGSAIARNVTVDPNPVPGDGYSHASQORVLSFNLGDMRPGDKKVFVTEFCPQRRGQIT 303

QY 301 NVATVTCGGHKCSANVTTVVNEPCQVQVNI SGADWSYVCKPVEYSISVSNPGLVLDVV 360
Db 304 NVATVTCGGHKCSANVTTVVNEPCQVQVNI SGADWSYVCKPVEYSISVSNPGLVLDVV 363
QY 361 IODTLPSTVLEAPGGEICCNKVVWRIKEMCPGETLQFKLVVKAQVPGRFTNOVAVTSE 420
Db 364 IODTLPSTVLEAPGGEICCNKVVWRIKEMCPGETLQFKLVVKAQVPGRFTNOVAVTSE 423
QY 421 SNCGTCTCAEATTHWKGLAATHMVCVLDNDPICVGENTVYRICVTVNRGSAEDTNVSLIL 480
Db 424 SNCGTCTCAEATTHWKGLAATHMVCVLDNDPICVGENTVYRICVTVNRGSAEDTNVSLIL 483
QY 481 KFSKELOPIASSGPTKGTISGNTVVPDALPKLGSKESEVFSVTLKGIAPGDARGEAILSS 540
Db 484 KFSKELOPIASSGPTKGTISGNTVVPDALPKLGSKESEVFSVTLKGIAPGDARGEAILSS 543
QY 541 DTLTSPVSDTENTHY 556
Db 544 DTLTSPVSDTENTHY 559

RESULT 5

US-09-438-185A-559
; Sequence 559, Application US/09438185A
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 559
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: CPh0557
US-09-438-185A-559

Query Match 100.0%; Score 556; DB 18; Length 559;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSLKIRRVTVLTALTSMSACFASGGTEAAVAESLITKIVASAETKPAVPMTAKKRVLR 60
Db 4 MSLKIRRVTVLTALTSMSACFASGGTEAAVAESLITKIVASAETKPAVPMTAKKRVLR 63
QY 61 RNKQPVQKSRGAFCDKEFYPCPEGRCQPVQAQOESCYGRLYSVKVNDCNVEICQSVPE 120
Db 64 RNKQPVQKSRGAFCDKEFYPCPEGRCQPVQAQOESCYGRLYSVKVNDCNVEICQSVPE 123
QY 121 YATVGSPIPIETILAIAGKDCVDVVIITQQLPCEAEFVSSDPETPTSDGKLWVKIDRLGAG 180
Db 124 YATVGSPIPIETILAIAGKDCVDVVIITQQLPCEAEFVSSDPETPTSDGKLWVKIDRLGAG 183
QY 181 DKCKITWWKPLKEGCGCFTAAATVCACPELRSYTKCGOPATCICKQEGPDCACLRCPVCYKI 240
Db 184 DKCKITWWKPLKEGCGCFTAAATVCACPELRSYTKCGOPATCICKQEGPDCACLRCPVCYKI 243
QY 241 EVVNTGSAIARNVTVDPNPVPGDGYSHASQORVLSFNLGDMRPGDKKVFVTEFCPQRRGQIT 300
Db 244 EVVNTGSAIARNVTVDPNPVPGDGYSHASQORVLSFNLGDMRPGDKKVFVTEFCPQRRGQIT 303

```
QY 301 NVATVTCGGHKCSANVTTVNPECVQVNISGADWSYCKPVEYSISVSNPGLVLDV 360
Db 304 NVATVTCGGHKCSANVTTVNPECVQVNISGADWSYCKPVEYSISVSNPGLVLDV 363
QY 361 IQDTLPSGVTVLEAPGGEICCNKVVRIKEMCPGETTQFKLVVRAQVPGRTNQAVTSE 420
Db 364 IQDTLPSGVTVLEAPGGEICCNKVVRIKEMCPGETTQFKLVVRAQVPGRTNQAVTSE 423
QY 421 SNGCTCTCAETTHHWKGLAATHMCLVDTNDPICVGVNTVYRICVTRNGSAEDTNVSLIL 480
Db 424 SNGCTCTCAETTHHWKGLAATHMCLVDTNDPICVGVNTVYRICVTRNGSAEDTNVSLIL 483
QY 481 KFSKELQPIASSGPTKGTISGNTVVFDPALPKLGSKEVSFEVTLKGIAPGDARGEAILSS 540
Db 484 KFSKELQPIASSGPTKGTISGNTVVFDPALPKLGSKEVSFEVTLKGIAPGDARGEAILSS 543
QY 541 DTLTSPVSDTENTHYV 556
Db 544 DTLTSPVSDTENTHYV 559

RESULT 6
US-10-020-269-38
; Sequence 38, Application US/10020269
; GENERAL INFORMATION:
; APPLICANT: Mitchell, William M.
; APPLICANT: Stratton, Charles W.
; TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE
; FILE REFERENCE: VDB98-01
; CURRENT APPLICATION NUMBER: US/10/020,269
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/025,596
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/023,921
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-08-14
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 38
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-020-269-38
```

```
Query Match 4.5%; Score 25; DB 24; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9e-18;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 RRKQPVEQKSRGAFCDKEFYPCEE 84
Db 1 RRKQPVEQKSRGAFCDKEFYPCEE 25
```

```
RESULT 7
US-09-841-132-441
; Sequence 441, Application US/09841132
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; PRIOR FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 441
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis serovar D
US-09-841-132-441
```

```
Query Match 4.5%; Score 25; DB 22; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.2e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 KITVWVKPLKEGCCFTAAATVCACPE 208
Db 181 KITVWVKPLKEGCCFTAAATVCACPE 205
```

```
RESULT 8
US-09-201-228A-1036
; Sequence 1036, Application US/09201228A
; GENERAL INFORMATION:
; APPLICANT: Griflais, Remy
; APPLICANT: Hoiseth, Susan K.
; APPLICANT: Zagursky, Robert John
; APPLICANT: Metcalf, Benjamin J.
; APPLICANT: Peek, Joel A.
; APPLICANT: Sankaran, Banumathi
; APPLICANT: Fletcher, Leah Diane
; TITLE OF INVENTION: CHLAMYDIA TRACHOMATIS GENOMIC SEQUENCE
; TITLE OF INVENTION: AND POLYPEPTIDES, FRAGMENTS THEREOF AND USES THEREOF, IN
; TITLE OF INVENTION: PARTICULAR FOR THE DIAGNOSIS, PREVENTION AND TREATMENT OF
; FILE REFERENCE: 9710-0004-999
; CURRENT APPLICATION NUMBER: US/09/201,228A
; PRIOR FILING DATE: 1998-11-30
; PRIOR APPLICATION NUMBER: US 60/107,077
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: FR 97-16034
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: FR 97-15041
; PRIOR FILING DATE: 1997-11-28
; NUMBER OF SEQ ID NOS: 5981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1036
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-09-201-228A-1036
```

```
Query Match 4.5%; Score 25; DB 16; Length 557;
Best Local Similarity 100.0%; Pred. No. 1.2e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 KITVWVKPLKEGCCFTAAATVCACPE 208
Db 181 KITVWVKPLKEGCCFTAAATVCACPE 205
```

```
RESULT 9
US-10-020-269-41
; Sequence 41, Application US/10020269
; GENERAL INFORMATION:
; APPLICANT: Mitchell, William M.
; APPLICANT: Stratton, Charles W.
; TITLE OF INVENTION: IDENTIFICATION OF ANTIGENIC PEPTIDE
; FILE REFERENCE: VDB98-01
; CURRENT APPLICATION NUMBER: US/10/020,269
; CURRENT FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/025,596
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/023,921
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-08-14
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 41
; LENGTH: 20
; TYPE: PRT
```

Query Match 2.9%; Score 16; DB 18; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;

Query Match 2.9%; Score 16; DB 18; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	25	GIEAAVAESLITKIVA	40
Db	1	GIEAAVAESLITKIVA	16

Search completed: May 25, 2002, 22:28:39
Job time: 222 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2002, 22:25:37 ; Search time 13.95 Seconds
(without alignments)
1128.285 Million cell updates/sec

Title: US-09-523-647-2
Perfect score: 556
Sequence: 1 MSKLIRRVTVLALTMASC.....ILSSDTLTSPVSDTENTHYV 556

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 116914 seqs, 28308587 residues

Word size : 12

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

- Database : Pending_Patents_AA_New.*
- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
 - 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
 - 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
 - 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
 - 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
 - 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
 - 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Prei. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
-----	-----	-----	-----	-----	-----

No matches found

Search completed: May 25, 2002, 22:29:01
Job time: 204 sec

